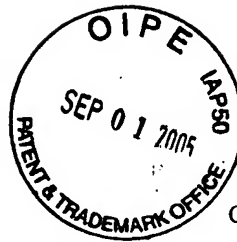


I hereby certify that this correspondence is being deposited with the U.S. Postal Service as Express Mail, Airbill No. EV644664535US, in an envelope addressed to: MS Amendment, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on the date shown below.

Dated: September 1, 2005

Signature: Enrique Moreno

(Enrique Moreno)



Docket No.: 393032040000
Client reference: Y1210U (H7927US)
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:

Jun ISHII et al.

Confirmation No.: 1959

Application No.: 10/642,542

Examiner: Jianchun Qin

Filed: August 14, 2003

Art Unit: 2837

For: SYNCHRONOUS PLAYBACK SYSTEM FOR
REPRODUCING MUSIC IN GOOD
ENSEMBLE AND RECORDER AND PLAYER
FOR THE ENSEMBLE

TRANSMITTAL OF ENGLISH TRANSLATIONS OF PRIORITY DOCUMENTS

MS Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In response to the Office Action dated June 3, 2005, enclosed herewith are English translations of the certified priority documents of the following Japanese Applications from which priority is claimed:

JP Application No. 2002-242481	Filed: August 22, 2002
JP Application No. 2002-242482	Filed: August 22, 2002
JP Application No. 2002-316806	Filed: October 30, 2002
JP Application No. 2002-316807	Filed: October 30, 2002

Dated: September 1, 2005

Respectfully submitted,

By Mehran Arjomand

Mehran Arjomand

Registration No.: 48,231
MORRISON & FOERSTER LLP
555 West Fifth Street, Suite 3500
Los Angeles, California 90013
(213) 892-5200

JAPAN PATENT OFFICE

This is to certify that the annexed is a true copy of the following application as filed with this Office.

Date of Application: August 22, 2002

Application Number: No. 2002-242481

[JP2002-242481]

Applicant(s): Yamaha Corporation

June 23, 2003

Commissioner,

Japan Patent Office Shinichiro OHTA (Seal)

Certification Number: Pat 2003-3048427

[Document Name]	Patent Application
[Reference Number]	C 30593
[Filing Date]	14th year of Heisei, August 22
[Acceptor]	To the Director General of the Patent Office
[International Patent Classification]	G11B 31 / 00
[Title of the Invention]	Apparatus for synchronized playback of audio data and performance data and method therefor
[Number of Claims]	7
[Inventor]	
[Address]	c/o YAMAHA CORPORATION, 10-1, Nakazawa-cho, Hamamatsu-shi, Shizuoka-ken
[Name]	Jun ISHII
[Inventor]	
[Address]	c/o YAMAHA CORPORATION, 10-1, Nakazawa-cho, Hamamatsu-shi, Shizuoka-ken
[Name]	Rei FURUKAWA
[Applicant for Patent]	
[Discriminative Number]	000004075
[Name]	YAMAHA CORPORATION
[Attorney]	
[Discriminative Number]	100098084
[Patent Attorney]	
[Name]	Kenji KAWASAKI
[Indication of Official Fee]	
[Ledger Code of Prepayment]	038265
[Amount of Money]	21000
[List of Documents Attached]	
[Document]	Specification 1
[Document]	Drawings 1
[Document]	Abstract 1

[DOCUMENT NAME]

Specification

[TITLE OF THE INVENTION]

Apparatus for synchronized playback of audio data and performance data and method therefor

[SCOPE OF THE PATENT CLAIM]

[Claim 1]

A recorder being characterized by having:

a first receiving means for receiving audio data representing audio waveform of a musical tune;

a second receiving means for receiving control data instructing performance control;

a generating means for generating reference data which abstracts the audio waveform represented by partial data of a portion of the aforesaid audio data; and

a recording means for recording the aforesaid reference data and for recording time data representing a time relation between a playback timing of the aforesaid partial data and a receiving timing of the aforesaid control data.

[Claim 2]

A player being characterized by having:

a first receiving means for receiving reference data which abstracts an audio waveform and performance data having control data for instructing the performance control and time data instructing an execution timing of control of said performance;

a second receiving means for receiving audio data representing an audio waveform of a musical tune;

a selecting means for selecting data representing audio waveform similar to an audio waveform represented by the aforesaid reference data as partial data from the

aforesaid audio data; and

a transmission means for transmitting the aforesaid control data at a timing determined by a playback timing of the aforesaid partial data and the aforesaid time data.

[Claim 3]

A recording method being characterized by having:

a first receiving step for receiving audio data representing audio waveform of a musical tune;

a second receiving step for receiving control data instructing performance control;

a generating step for generating reference data which abstracts the audio waveform represented by partial data of a portion of the aforesaid audio data; and

a recording step for recording the aforesaid reference data and for recording time data representing a time relation between a playback timing of the aforesaid partial data and a receiving timing of the aforesaid control data.

[Claim 4]

A playback method being characterized by having:

a first receiving step for receiving reference data which abstracts an audio waveform and performance data having control data for instructing the performance control and time data instructing an execution timing of control of said performance;

a second receiving step for receiving audio data representing an audio waveform of a musical tune;

a selecting step for selecting data representing audio waveform similar to an audio waveform represented by the aforesaid reference data as partial data from the aforesaid audio data; and

a transmission step for transmitting the aforesaid control data at a timing

determined by a playback timing of the aforesaid partial data and the aforesaid time data.

[Claim 5]

A program to have a computer execute:

a first receiving step for receiving audio data representing audio waveform of a musical tune;

a second receiving step for receiving control data instructing performance control;

a generating step for generating reference data which abstracts the audio waveform represented by partial data of a portion of the aforesaid audio data; and

a recording step for recording the aforesaid reference data and for recording time data representing a time relation between a playback timing of the aforesaid partial data and a receiving timing of the aforesaid control data.

[Claim 6]

A program to have a computer execute:

a first receiving step for receiving reference data which abstracts an audio waveform and performance data having control data for instructing the performance control and time data instructing an execution timing of control of said performance;

a second receiving step for receiving audio data representing an audio waveform of a musical tune;

a selecting step for selecting data representing audio waveform similar to an audio waveform represented by the aforesaid reference data as partial data from the aforesaid audio data; and

a transmission step for transmitting the aforesaid control data at a timing determined by a playback timing of the aforesaid partial data and the aforesaid time data.

[Claim 7]

A recording medium recording reference data, which abstracts audio waveform, and performance data having control data instructing performance control and time data instructing an execution timing of control of said performance.

[DETAILED EXPLANATION OF THE INVENTION]

[0001]

[TECHNICAL FIELD OF THE INVENTION]

The present invention relates to an apparatus and a method for playing back performance data including information relating to performance control of a musical tune in synchronization with playback of audio data.

[0002]

[PRIOR ART]

There is an apparatus for reading out audio data from a recording medium such as a music CD (Compact Disc) and generating sounds from the readout audio data to be output as a means for playing back a musical tune. There is an automatic performance apparatus for reading out data including information on performance control of a musical tune from a recording medium such as an FD (Floppy Disk) and for controlling tone generation of a tone generator by using the readout data as another means for playing back the musical tune. There is MIDI data created by complying with the MIDI (Musical Instrument Digital Interface) standard as the data including information relating performance control of the musical tune.

[0003]

Recently, a method for synchronizing the automatic performance with MIDI data with the playback of audio data recorded on the music CD has been proposed. There is a method of using time codes recorded on the music CD as one of them (see patent

document 1 and patent document 2, for example). The methods will be explained below.

[0004]

First, the audio data in the music CD and time codes are reproduced by a music CD player. Then, the audio data is output as a sound and the time codes are supplied to a recorder. Herein, the time code is data associated with the certain amount of audio data and each time code represents a lapse of time from the a start of the musical tune to a playback timing of the audio data associated with said time code. The musical instrument is played with the playback of the music CD and the MIDI data are sequentially supplied from the musical instrument to the recorder. The recorder receives the MIDI data from the musical instrument and records the MIDI data on the recording medium with the time information representing the receiving timing. The recorder receives the time code from the music CD player and records this in the recording medium with the time information representing receiving timing. As a result, a file in which the time code and the MIDI data are mixed is created in the recording medium. In this file, respective time codes and the MIDI data have time information representing lapses of time from the musical tune playback starting timing until respective playback timing.

[0005]

Thus, the audio data of the same musical tune is played back from a music CD after the MIDI data and the time code are recorded on the recording medium, the MIDI data is read out from the recording medium in synchronization therewith and the automatic performance is realized. The operations are as follows.

[0006]

First, the audio data and the time code are reproduced from the music CD by a music CD player. Then, the audio data is output as sound and the time code is supplied to

the player of the MIDI data. The player reads out the MIDI data stored in the file based on the time information recorded therewith and sequentially transmits those to a musical instrument capable of automatic performance with the MIDI data. At that time, the player adjusts the time difference between the playback of the audio data of the music CD and the playback of the MIDI data based on the time code received from the CD player and the time code read out from the file together with the MIDI data. As a result, the synchronized playback of the audio data of the music CD and the MIDI data are realized.

[0007]

[Patent Publication 1]:

Patent Application No. 2002 - 7872

[Patent Publication 2]:

Patent Application No. 2002 - 7873

[0008]

[PROBLEM TO BE SOLVED BY THE INVENTION]

However, the synchronized playback of the audio data of the music CD and the MIDI data is not possible by a method of using the time codes of the music CD for a music CDs labeled with different time codes for the same musical tune.

[0009]

Currently, there are different versions of music CDs for a single musical tune. Though the content of the musical tune itself is the same, silent time periods at the start of the musical tune differ among the music CD and as a result, the time codes when the performance of the musical tunes actually starts are different very much. In other words, when the MIDI data for synchronized performance, which is created by the technology using the conventional time code, is used for the music CD of the same musical tune

having the different version, the performance by the MIDI data starts before the actual performance of the musical tune starts, or the performance by the MIDI data does not start for a while after the performance of the musical tune starts; the performance by the MIDI data is shifted from the musical tune of the music CD.

[0010]

Accordingly, it encounters a problem that different versions of MIDI data for synchronized performance have to be prepared depending on variations of the time codes corresponding to the playback start timings of the actual musical tunes for the music CDs recording the audio data for the same musical tune by using the conventional time code technology.

[0011]

By contemplating the above mentioned circumstances, it is an object of the present invention to provide a recorder, a player, a recording method, a playback method and program which plays back the performance data such as the MIDI data synchronously with plural versions of audio data having different playback start timings of the actual musical tune for the audio data for the same musical tune.

[0012]

[MEANS TO SOLVE THE PROBLEM]

To solve the above explained problem, the present invention provides a recorder being characterized by having:

a first receiving means for receiving audio data representing audio waveform of a musical tune;

a second receiving means for receiving control data instructing performance control;

a generating means for generating reference data which abstracts the audio waveform represented by partial data of a portion of the aforesaid audio data; and

a recording means for recording the aforesaid reference data and for recording time data representing a time relation between a playback timing of the aforesaid partial data and a receiving timing of the aforesaid control data.

[0013]

The present invention provides a player being characterized by having:

a first receiving means for receiving reference data which abstracts an audio waveform and performance data having control data for instructing the performance control and time data instructing an execution timing of control of said performance;

a second receiving means for receiving audio data representing an audio waveform of a musical tune;

a selecting means for selecting data representing audio waveform similar to an audio waveform represented by the aforesaid reference data as partial data from the aforesaid audio data; and

a transmission means for transmitting the aforesaid control data at a timing determined by a playback timing of the aforesaid partial data and the aforesaid time data.

[0014]

The present invention provides a recording method being characterized by having:

a first receiving step for receiving audio data representing audio waveform of a musical tune;

a second receiving step for receiving control data instructing performance control;

a generating step for generating reference data which abstracts the audio waveform represented by partial data of a portion of the aforesaid audio data; and

a recording step for recording the aforesaid reference data and for recording time data representing a time relation between a playback timing of the aforesaid partial data and a receiving timing of the aforesaid control data.

[0015]

The present invention provides a playback method being characterized by having:

a first receiving step for receiving reference data which abstracts an audio waveform and performance data having control data for instructing the performance control and time data instructing an execution timing of control of said performance;

a second receiving step for receiving audio data representing an audio waveform of a musical tune;

a selecting step for selecting data representing audio waveform similar to an audio waveform represented by the aforesaid reference data as partial data from the aforesaid audio data; and

a transmission step for transmitting the aforesaid control data at a timing determined by a playback timing of the aforesaid partial data and the aforesaid time data.

[0016]

The present invention provides a program to have a computer execute these recording method and playback method.

[0017]

The present invention provides a recording medium recording reference data, which abstracts audio waveform, and performance data having control data instructing performance control and time data instructing an execution timing of control of said performance.

[0018]

Using the apparatus, method, program and recording medium implemented by these configurations, the position of the reference data of the audio data along the time axis is determined based on the similarity of the waveforms represented by the audio data upon playback of the audio data and the playback timing of the control data is determined based on the position of the reference data along the time axis. As a result, the audio data and the control data are synchronously reproduced.

[0019]

The recorder implemented by the present invention may have a third receiving means for receiving a time code representing the playback timing of the aforesaid audio data and the aforesaid recording means may generate the aforesaid time data based on the time information represented by the aforesaid time code.

The player implemented by the present invention may have a third receiving means for receiving the time code representing the playback timing of the aforesaid audio data and the aforesaid transmission means may transmit the aforesaid control data based on the time information represented by the aforesaid time code.

[0020]

Using the recorder and the player having said configuration, since the time is measured for the audio data played by a player at a biased playback speed with the time codes, the control data is correctly and synchronously played.

[0021]

In the recorder implemented by the present invention, the aforesaid generating means may have a filter means for eliminating DC components of the audio waveform represented by the input data.

In the recorder implemented by the present invention, the aforesaid generating

means may have a filter means for extracting a specific frequency band included in the audio waveform represented by the input data.

[0022]

In the player implemented by the present invention, the aforesaid selecting means may have a generating means for generating discrimination data which abstracts the audio waveform represented by a part of the aforesaid audio data, and the aforesaid generating means may have a filter means for eliminating the DC components of the audio waveform represented by the input data.

In the player implemented by the present invention, the aforesaid selecting means may have a generating means for generating discrimination data which abstracts the audio waveform represented by a part of the aforesaid audio data, and the aforesaid generating means may have a filter means for extracting a specific frequency band included in the audio waveform represented by the input data.

[0023]

By using the recorder and the player with said configurations, when the position of the reference data along the time axis with respect to the audio data is determined based on the similarity of the audio waveforms represented by the audio data, the position is determined with a high accuracy.

[0024]

In the recorder implemented by the present invention, the aforesaid generating means may have a down sampling means for sampling down the input data.

In the player implemented by the present invention, the aforesaid selecting means may have a generating means for generating discrimination data which abstracts the audio waveform represented by a part of the aforesaid audio data, and the aforesaid generating

means may have a down sampling means for sampling down the input data.

[0025]

By using recorder and the player with said configurations, the data amount of the reference data is made small, and the recording, transmission and reception of data are made easy.

[0026]

In the player implemented by the present invention, the aforesaid selecting means may have a generating means for generating discrimination data which abstracts the audio waveform represented by a part of the aforesaid audio data and selects the aforesaid partial data based on the index acquired by dividing the sum of products of the aforesaid reference data and the aforesaid discrimination data by the sum of the squares of the aforesaid reference data.

[0027]

In the player implemented by the present invention, the aforesaid selecting means may have a generating means for generating discrimination data which abstracts the audio waveform represented by a part of the aforesaid audio data and selects the aforesaid partial data based on the index acquired by dividing a square of the sum of products of the aforesaid reference data and the aforesaid discrimination data by the product of the sum of the squares of the aforesaid reference data and the sum of the squares of the aforesaid reference data.

[0028]

In the player implemented by the present invention, the aforesaid selecting means may have a generating means for generating discrimination data which abstracts the audio waveform represented by a part of the aforesaid audio data and selects the aforesaid partial

data based on a variation rate of the sum of products of the aforesaid reference data and the aforesaid discrimination data.

[0029]

Using the player with said configuration, when the position of the reference data along the time axis with respect to the audio data is determined based on the similarity of the audio waveform represented by the audio data, the position is determined with high accuracy.

[0030]

[EMBODIMENT OF THE INVENTION]

[1] First embodiment

[1.1] Structure, function and data format

[1.1.1] Whole configuration

Fig. 1 is a view to show a configuration of a synchronized recorder and player SS implemented by a first embodiment of the present invention. The synchronized recorder and player SS comprises a music CD drive 1, an FD drive 2, an automatic player piano 3, a tone generating portion 4, a manipulating display 5, and a controller 6.

[0031]

The music CD drive 1, FD drive 2, automatic player piano 3, tone generating portion 4 and manipulating display 5 are connected to the controller 6 by communication lines, respectively. The automatic player piano 3 and the tone generating portion 4 are directly connected each other by the communication line.

[0032]

[1.1.2] Music CD drive

The audio data stored in the music CD includes audio data representing audio

information and time codes representing playback timings of the audio data. The music CD drive 1 reads out the audio data from the loaded music CD under instructions from the controller 6 and sequentially outputs the audio data included in the audio data. The music CD drive 1 is connected to a communication interface 65 in the controller 6 by a communication line.

[0033]

The audio data, which is output from the music CD drive 1, is 16 bit digital audio data in two channels in left and right quantized at a sampling frequency of 44,100 Hz in 16 bits. The data output from the music CD drive 1 does not include the time code. Since the configuration of the music CD drive 1 is similar to a general music CD drive which is capable of outputting the digital audio data, the explanation will be omitted.

[0034]

[1.1.3] FD drive

The FD drive 2 records SMF (Standard MIDI File) in the FD or reads out the SMF recorded in the FD in order to transmit the readout SMF. The FD drive 2 is connected to the communication interface 65 in the controller 6 by the communication line. Since the configuration of the FD drive 2 is similar to a general FD drive, the explanation will be omitted.

[0035]

[1.1.4] MIDI event and SMF

The SMF is a file including the MIDI event serving as the performance control data complied with the MIDI standard and delta time serving as data representing the execution timing of the respective MIDI events. The MIDI events and the format of the SMF will be explained with reference to Fig. 2 and Fig. 3.

[0036]

A note-on event, a note-off event and a system exclusive event are shown in Fig. 2 as examples of the MIDI event. The note-on event is a MIDI event which instructs the tone generation of a musical tone and comprises 9nH (n represents a channel number, H represents a hexadecimal number, and which will be the same hereinafter) representing the tone generation, a note number representing a pitch, and a velocity representing the strength of the tone generation (or a velocity of striking a key). Similarly, the note-off event is a MIDI event which instructs the tone extinction of a musical tone and comprises 8nH representing the tone extinction, a note number representing a pitch, and a velocity representing the strength of the tone extinction (or a velocity of releasing a key). On the other hand, the system exclusive event is a MIDI event which transmits, receives or records data formatted by a product or software manufacturer's discretion, and comprises F0H representing a start of the system exclusive event, a data length, data and F7H representing an end of the system exclusive event. In such a way, the MIDI event does not include time information and is used for tone generation, tone extinction of the musical tone and other control in real time.

[0037]

Fig. 3 shows the general overview of the SMF format. The SMF includes a header chunk and a track chunk. The header chunk includes control data relating data format and time unit information included in the track chunk. The track chunk includes the MIDI event and delta time representing an execution timing of respective MIDI events.

[0038]

The SMF has expressions for delta time, namely, the one is expressed in time called as a clock representing the relative time with respect to a MIDI event immediately

before and the other is expressed by a combination time of hour, minute, second and frame to represent the absolute time from the head of the musical tune. To make the explanation easy, the delta time is defined as the absolute time from the base time and is expressed in second in the explanation below.

Moreover, the MIDI data is the comprehensive name for data created by complying with the MIDI standard in the present specification.

[0039]

[1.1.5] Automatic player piano

The automatic player piano 3 is a musical tone generator which outputs an acoustic piano tone and an electronically synthesized piano tone in response to a key manipulation and a pedal manipulation by the user of the synchronized recorder and player SS. The automatic player piano 3 generates a MIDI event in response to the key manipulation and the pedal manipulation by the user and transmits the generated MIDI event. Further, the automatic player piano 3 receives the MIDI event and automatically plays with acoustic piano sounds and electronically synthesized piano tones in response to the received MIDI events.

[0040]

The automatic player piano 3 comprises a piano 31, key sensors 32, pedal sensors 33, a MIDI event control circuit 34, a tone generator 35 and a driving part 36.

[0041]

The key sensor 32 and the pedal sensor 33 are provided on each of the plural keys and the plural pedals of the piano 31 in order to detect the positions of keys and pedals, respectively. The key sensor 32 and pedal sensor 33 transmit the detected position information, identification number corresponding to each of the keys and pedals,

respectively, and detected time information to the MIDI event control circuit 34.

[0042]

The MIDI event control circuit 34 receives the position information of the keys and pedals, respectively, from the key sensors 32 and the pedal sensors 33 together with the identification information of the keys and pedals and the time information, and immediately generates MIDI events such as a note-on event, note-off event or the like from the information in order to output the generated MIDI event to the controller 6 and the tone generator 35. The MIDI event control circuit 34 receives the MIDI event from the controller 6 and transfers the received MIDI event to the tone generator 35 or the driving part 36. Further, the MIDI event control circuit 34 is under the instruction of the controller 6 to determine which the MIDI event received from the controller 6 is transferred to the tone generator 35 or the driving part 36.

[0043]

The tone generator 35 receives the MIDI events from the MIDI event control circuit 34 and outputs the sound information of various musical instruments as digital audio data in the left and right channels based on the received MIDI events. The tone generator 35 electronically synthesizes the digital audio data at a pitch designated by the received MIDI event and transmits it to a mixer 41 in the tone generating portion 4.

[0044]

The driving parts 36 are provided on the respective keys and pedals of the piano 31 and comprises a group of solenoids for driving these and a control circuit for controlling the group of solenoids. When the control circuit of the driving parts 36 receives the MIDI events from the MIDI event control circuit, it adjusts current to be supplied to the solenoid provided on a corresponding key or pedal in order to adjust magnetic flux generated by the

solenoid and the key or pedal is operated in response to the MIDI event.

[0045]

[1.1.6] Tone generator

The tone generating portion 4 receives the audio data from the automatic player piano 3 and the controller 6 and converts the received audio data into sounds to be output. The tone generating portion 4 comprises the mixer 41, a D/A converter 42, an amplifier 43 and a speaker 44.

[0046]

The mixer 41 is a digital stereo mixer which receives plural sets of digital audio data in the two channels, left and right, and converts these into a pair of left and right digital audio data. The mixer 41 receives the digital audio data from the tone generator 35 of the automatic player piano 3 and at the same time, receives the digital audio data, which is read out by the music CD drive 1 from the music CD, through the controller 6. The mixer 41 calculates an average of the received digital audio data and transmits this to the D/A converter 42 as a pair of digital audio data in right and left.

[0047]

The D/A converter 42 receives the digital audio data from the mixer 41 and converts the received digital audio data into the analog audio signal to be output to the amplifier 43. The amplifier 43 amplifies the analog audio signal, which is input from the D/A converter 42, and outputs it to the speaker 44. The speaker 44 converts the analog audio signal, which is input from and amplified by the amplifier 43, into sounds. As a result, the audio data recorded in the music CD and the audio data generated by the tone generator 35 are output from the tone generating portion 4 as stereo sounds.

[0048]

[1.1.7] Manipulation display

The manipulation display 5 is a user interface when a user of the synchronized recorder and player SS manipulates the synchronized recorder and player SS.

[0049]

The manipulation display 5 includes key pads when the user depresses to give instructions to the synchronized recorder and player SS and a liquid crystal display for confirming the state of the synchronized recorder and player SS. When the key pad is depressed by the user, the manipulation display 5 outputs a signal corresponding to the depressed key pad to the controller 6. The manipulation display 5 receives bit map data including information of characters and figures and displays the characters and figures based on the received bit map data on the liquid crystal display.

[0050]

[1.1.8] Controller

The controller 6 controls the entire synchronized recorder and player SS. The controller 6 comprises a ROM (Read Only Memory) 61, a CPU (Central Processing Unit) 62, a DSP (Digital Signal Processor) 63, a RAM (Random Access Memory) 64 and a communication interface 65. The components are mutually connected each other through a bus.

[0051]

The ROM 61 is a non-volatile memory for storing various kinds of control program. The control program, which is stored in the ROM 61, includes program for general control routine and program which causes the CPU 62 to execute routines for recording operations and playback operations of the SMF which will be mentioned later. The CPU 62 is a microprocessor, which executes general purpose processings, and reads

out the control program from the ROM 61 and executes the control routines in accordance with the readout control program. The DSP 63 is a microprocessor, which processes the digital audio data at a high speed, executes data generation routine for correlation discrimination and filter operation necessary for the correlation discrimination routine under the control of the CPU 2, which will be mentioned later, on the digital audio data received from the music CD drive 1 and the FD drive 2 by the controller 6, and transmits the resulting data to the CPU 62. The RAM 64 is a volatile memory and temporarily stores the data used by the CPU 62 and DSP 63. The communication interface 65 is an interface which is capable of transmitting and receiving the digital data in various formats, converts the format necessary for digital data transmitted or received between the music CD drive 1, FD drive 2, automatic player piano 3, tone generating portion 4, and manipulation display 5, and relays the data between the respective devices and the controller 6.

[0052]

[1.2] Operation

Next, operations of the synchronized recorder and player SS will be explained.

[1.2.1] Recording operation

First, operations of the synchronized recorder and player SS, when a user of the synchronized recorder and player SS plays a piano in synchronization with the playback of a commercially available music CD and the information of the performance is recorded on an FD as the MIDI data, will be explained. The music CD used during the recording operation, which will be explained below, is called as a music CD-A in order to discriminate the music CD used during the playback operation mentioned later.

[0053]

[1.2.1.1] Start operation of recording

The user sets the music CD-A in the music CD drive 1 and an empty FD in the FD drive 2. Subsequently, the user depresses the key pads of the manipulation display 5 corresponding to the recording start of the performance data. The manipulation display 5 outputs the signal corresponding to the depressed key pad to the controller 6.

[0054]

The CPU 62 of the controller 6 receives the signal corresponding to the recording start of the performance data from the manipulation display 5 and transmits a playback instruction of the music CD to the music CD drive 1. In response to the playback instruction, the music CD drive 1 sequentially transmits the audio data recorded in the music CD-A to the controller 6. The controller 6 receives the data for a pair of right and left channels for every $1/44100$ second from the music CD drive 1. Hereunder, the data values for a pair of the right and left channels are expressed as $(R(n), L(n))$, and the pair of data values or respective sets of data values generated from the pair of data values in the data generation routine for correlation discrimination are called as "sample values". The $R(n)$ and $L(n)$ represent data values in the right channel and the left channel, respectively, and they are either of integers ranging from - 32768 to 32767. n is an integer representing an order of the audio data and increases from the start of the data such as 0, 1, 2

[0055]

[1.2.1.2] Transmission of the audio data to the tone generator

First, the CPU 62 receives the sample values, namely, $(R(0), L(0))$, $(R(1), L(1))$, $(R(2), L(2))$..., and transmits the received sample values to the tone generating portion 4. The tone generating portion 4 receives the sample values from the controller 6 and converts them to the sounds to be output. As a result, the user listens to a musical tune

recorded in a music CD-A.

[0056]

[1.2.1.3] Recording raw reference audio data into RAM

The CPU 62 transmits the received sample values to the tone generating portion 4 and records the sample values corresponding to a certain period of time at the head of the musical tune in the received sample values into the RAM 64.

In the present embodiment, the CPU 62 records 2^{16} pairs, namely, 65536 pairs of sample values into the RAM 64, as an example. Further, 65536 sample values cover data for about 1.49 seconds.

[0057]

Then, the CPU 62 judges whether the absolute value of each of sample values exceeds a previously defined threshold or not, with respect to each of the sample values. Specifically, it is assumed that the threshold value is 1000 and the CPU 62 gives an affirmative result in the judgment when either of the absolute values of $R(n)$ or $L(n)$ is larger than 1000 after comparison.

[0058]

Hereunder, as an example for explanation, it is assumed that at 52156^{th} pairs of sample values, namely, at $(R(52156), L(52156))$, the absolute values of $R(52156)$ or $L(52156)$ exceeds a predetermined threshold for the first time with respect to the audio data of the music CD-A. Accordingly, the CPU 62 gets negative results for $(R(0), L(0))$ to $(R(52155), L(52155))$ through comparing discrimination. During this, the CPU 62 does not store these sample values in the RAM 64. As a result, sample values for the silent or substantially silent part included at the head of the musical tune are not recorded in the RAM 64. In this case, the playing time for the sample values at the head which is

not stored is about 1.18 seconds.

[0059]

Thereafter, the CPU 62 receives (R(52156), L(52156)) and gets an affirmative result after comparing the sample values therewith. When acquiring the affirmative result after the comparison, the CPU 62 stores 65536 pairs of samples thereafter, namely, (R(52156), L(52156)) to (R(117691), L(117691)) in the RAM 64. Hereunder, a series of sample values is called as "raw reference audio data".

[0060]

[1.2.1.4] Start of measuring

The CPU 62 receives the last sample value of the raw reference audio data, namely, (R(117691), L(117691)), and finishes recording the raw reference audio data, and starts measuring time from the timing.

[0061]

[1.2.1.5] Generation of processed reference audio data

The CPU 62 finishes the recording of the raw reference audio data and sends an instruction, which executes the data generation process for correlation discrimination on the raw reference audio data, to a DSP 63. The data generation process for correlation discrimination is a process for generating audio data sampled at a sampling frequency of about 172.27Hz for correlation discrimination process from the audio data sampled at a sampling frequency of 44,100Hz. The correlation discrimination process is a process to judge similarity of two pairs of audio data and the details will be mentioned later. Hereunder, the data generation process for correlation discrimination will be explained with reference to Fig. 4.

[0062]

The DSP 63 receives the instruction to execute the data generation process for correlation discrimination on the raw reference audio data from the CPU 62 and reads out the raw reference audio data stored in the RAM 64 (step S1). Subsequently, the DSP 63 calculates an arithmetic average of the left and right values to the respective sample values of the raw reference audio data and converts the stereo data into a monaural data (step S2). The conversion process into the monaural is a process to reduce the workload on the DSP 63 in processes after this step.

[0063]

Subsequently, the DSP 63 puts a series of sample values converted into the monaural signal in a high pass filtering (step S3). The DC components in the audio waveform represented by the series of sample values are eliminated by this high pass filtering and the sample values are uniformly distributed in positive and negative sides. Two pairs of audio data are compared and discriminated based on cross correlation values in the correlation discrimination process, and the preciseness of discrimination is enhanced if the sample values are uniformly distributed on positive and negative sides when the cross correlation values are compared. In other words, the process in this step is a process to improve the accuracy of the judgment in the correlation discrimination process.

[0064]

Subsequently, the DSP 63 calculates absolute values of respective sample values after the high pass filtering (step S4). The process in the step calculates substitute values of power of the respective samples. Since the absolute values have smaller values than square values representing the power and are easily processed, the present embodiment uses the absolute values instead of the square values of respective sample values. The square values may be calculated instead of the absolute values of the respective sample

values in this step when the performance of the DSP 63 is high.

[0065]

Subsequently, the DSP 63 filters the series of sample values, which are converted into the absolute values in the step S4, through a comb filter (step S5). The process in this step extracts the low frequency components, of which the variation in the waveform is easily to be detected, from the audio signal waveform represented by the series of sample values. A low pass filter is normally used to extract the low frequency components; since the comb filter applies less load to the DSP 63 than the low pass filter, the comb filter is replaced with the low pass filter in the present embodiment.

[0066]

Fig. 5 shows a configuration of an example of the comb filter to be employed in the step S5. In Fig. 5, a process represented by a square rectangular means a delay process and k in z^{-k} means that the delay time in the delay process is (sampling cycle $\times k$). As mentioned previously, the sampling frequency of the music CD is 44100 Hz and the sampling period is $1 / 44100$ second. On the other hand, the process represented by a triangle means a multiplication and a value indicated in the triangle means a coefficient of the multiplication. In Fig. 5, K is expressed by a following expression (1).

[Expression 1]

$$K = \frac{44100 - \pi \times f}{44100 + \pi \times f} \quad \dots\dots (1)$$

[0067]

The multiplication using K as a coefficient gives the comb filter a function of a high pass filter having a cutoff frequency f . As a result, the DC components in the audio waveform represented by the series of sample values are eliminated again by the filtering process in this step. Moreover, the values of k and f are arbitrarily varied and are

empirically calculated in order to enhance the accuracy of discrimination in the correlation discrimination process.

[0068]

Subsequently, the DSP 63 filters the series of sample values, which are filtered in the step S5, through a low pass filter (step S6). The process in this step avoids aliasing noise in a down sampling process rendered in a next step S7. Since the data at the sampling frequency of 44100 Hz are sampled down to a sampling frequency of about 172.27 Hz, the frequency components of about 86.13 Hz, which is a half thereof, or higher need to be eliminated in order to avoid the aliasing noise. However, the high frequency components are not sufficiently eliminated in the filtering process in the step S5 using the comb filter due to the characteristics of the comb filter. Accordingly, the remaining frequency components of about 86.13 Hz or higher are eliminated by the filtering process using the low pass filter in this step. If the performance of the DSP 63 is high, a filtering process using a single low pass filter with a high accuracy is acceptable instead of the filtering process using two filters in the step S5 and step S6.

[0069]

Subsequently, the DSP 63 samples down the series of sample values filtered in the step S6 by 1/256 (step S7). In other words, the DSP 63 extracts one sample value from every 256 sample values. As a result, the number of the series of sample data is reduced from 65536 to 256. Hereunder, each of the sample values acquired from the process in the step S7 are expressed by $X(m)$. However, m is an integer ranging from 0 to 255. The series of sample values, namely, $X(0)$ to $X(255)$ is called as "processed reference audio data". The DSP 63 stores the processed reference audio data in the RAM 64 (step S8).

[0070]

[1.2.1.6] Recording MIDI event into RAM

The DSP 63 generates the processed reference audio data as mentioned the above and the user starts the performance by the piano 31. In other words, the CPU 62 finishes the recording of the raw reference audio data and starts time measuring; while listening to a musical tune from the music CD-A, which is output from the tone generating portion 4, the user depresses keys and manipulates the pedal of the piano 31 together with the musical tune.

The motions of the keys and the pedals are detected as the performance information by using the piano 31 of the user through the key sensors 32 and the pedal sensors 33 and are converted into MIDI events by a MIDI event control circuit 34 to be transmitted to the controller 6.

[0071]

The CPU 62 in the controller 6 receives the MIDI events from the automatic player piano 3 and records a measured time value upon receiving the MIDI event, namely, a delta time representing a lapse of time from a timing when the CPU 62 receives the last sample value of the raw reference audio data by a timing when the MIDI event is received, in the RAM 64 with the MIDI event. Fig. 6 is an illustrative view representing the relation with respect to time between the audio of the music CD-A and MIDI events. In Fig. 6, the CPU 62 starts the time measurement after about 2.67 seconds after starting the play of the audio data in the music CD-A and a first MIDI event, a second MIDI event and a third MIDI event are received by the CPU 62 after 1.25 seconds, 2.63 seconds and 3.71 seconds, respectively, from the timing.

[0072]

[1.2.1.7] Recording SMF in FD

After finishing playing the musical tune in the music CD-A and the performance by the user with the piano 31, the user depresses a key pad on the manipulation display 5 corresponding to an end of recording of the performance data. The manipulation display 5 transmits a signal corresponding to the depressed key pad to the controller 6. The CPU 62 receives a signal representing the end of recording the performance data from the manipulation display 5 and transmits an instruction for stopping the play of the music CD to the music CD drive 1. The music CD drive 1 stops playing the music CD-A in response to the instruction for stopping the play.

[0073]

Subsequently, the CPU 62 reads out the processed reference audio data generated by the DSP 63, the MIDI events and the delta time generated through performance by the user by a piano 31 from the RAM 64. The CPU 62 combines these readout data and forms a track chunk of the SMF. The CPU 62 attaches a header chunk corresponding to the created track chunk and forms the SMF.

[0074]

Fig. 7 is a view to show the overview of the SMF created by the CPU 62. A system exclusive event including the processed reference audio data is recorded together with the delta time therefor in the header of the data area in the track chunk. The delta time is 0.00 second. Following the system exclusive event including the processed reference audio data, MIDI events associated with the performance by the user with the piano 31 are sequentially recorded. In the example in Fig. 6, a first MIDI event is a note-on event at C5; a second MIDI event is a note-on event at E6; and a third MIDI event is a note-off event at C5 by the performance of the user, and the delta time corresponding to

these are 1.25 seconds, 2.63 seconds and 3.71 seconds, respectively.

[0075]

The CPU 62 completes the generation of the SMF and transmits the generated SMF to the FD drive 2 together with a write instruction. The FD drive 2 receives the write instruction and the SMF from the CPU 62 and writes the SMF into a loaded FD.

[0076]

Fig. 6 shows the time relation between the audio data in the music CD-A and MIDI events to be written in the SMF. In the following explanation, a timing from a playback start timing of the music CD-A, which is 0 second, is labeled with (T) as a suffix and a delta time in the SMF is labeled with (D) as a suffix in order to discriminate two different timings.

[0077]

First, the absolute value of the audio data in the music CD-A becomes larger than a threshold value of 1000 at about 1.18 second (T) and the raw reference audio data is started to be recorded. Thereafter, the raw reference audio data is recorded for about 1.49 seconds, namely, until 2.67 seconds (T).

[0078]

Subsequently, the time measurement has been started from a timing of about 2.67 seconds (T) as 0 second in order to calculate the delta time. Thereafter, a first event is generated at 1.25 second (D), namely, about a timing of about 3.92 seconds (T) and the event is recorded. As similarly, a second event is generated at 2.63 second (D), namely, about 5.30 seconds (T) and a third event is generated at 3.71 seconds (D), namely, about 6.38 seconds (T) and these events are recorded.

[0079]

Further, as shown in a lower row in Fig. 7, the playback time of the raw reference audio data corresponding to the processed reference audio data is before 0.00 second (D) and the processed reference audio data in the SMF is recorded at an area of 0.00 second (D) as the system exclusive data.

[0080]

[1.2.2] Playback operation

Subsequently, the operations to play back the SMF recorded by the above mentioned method and to synchronize the audio data of the music CD with the MIDI data of the SMF will be explained. The music CD used during the playback operation includes the musical tune same as those of the music CD-A used in the above mentioned recording operation, however, its version is different and a time period from a playback start of the music CD to the start of the musical tune and the level of the audio waveform representing the audio data are different. Further, since audio effects of the audio data of the music CD are edited when data for press is created from the master data of the musical tune, the contents are slightly different from the same musical tune data in the music CD-A. Accordingly, the music CD used in the playback operation, which will be explained hereunder, is called as a music CD-B in order to discriminate it from the music CD-A.

[0081]

[1.2.2.1] Playback start manipulation

The user loads a music CD-B on the music CD drive 1 and an FD, on which the SMF is recorded, on the FD drive 2. Subsequently, the user depresses the key pad of the manipulation display 5 corresponding to the playback start of the performance data. The manipulation display 5 outputs a signal corresponding to the depressed key pad to the controller 6.

[0082]

The CPU 62 receives a signal instructing the playback start of the performance data from the manipulation display 5 and transmits a transmission instruction of the SMF to the FD drive 2. The FD drive 2 reads out the SMF from the FD in response to the transmission instruction of the SMF and transmits the readout SMF to the controller 6. The CPU 62 receives the SMF from the FD drive 2 and stores the received SMF in the RAM 64.

[0083]

Subsequently, the CPU 62 transmits a playback instruction of the music CD to the music CD drive 1. The music CD drive 1 sequentially transmits the audio data stored in the music CD-B to the controller 6 in response to the playback instruction. The controller 6 receives a pair of data in the left and right channels from the music CD drive 1 for every $1/44100$ second. Herein, the data values received from the music CD drive 1 by the CPU 62 are represented by $(r(n), l(n))$. The ranges of the values $r(n)$ and $l(n)$ and the definitions of n and "sample values" used hereunder are similar to those of $R(n)$ and $L(n)$.

[0084]

[1.2.2.2] Transmission of audio data to tone generator

The CPU 62 receives the sample values, namely, $(r(0), l(0))$, $(r(1), l(1))$, $(r(2), l(2))$, ... from the music CD drive 1 and transmits the received sample values to the tone generating portion 4. The tone generating portion 4 receives the sample values from the controller 6 and converts it to sounds to be output. As a result, the user can listen to a musical tune recorded on the music CD-B.

[0085]

[1.2.2.3] Correlation discrimination process

The CPU 62 transmits the sample values received from the music CD drive 1 to the tone generating portion 4 and, at the same time, transmits an execution instruction of the correlation discrimination process to the DSP 63 and sequentially transmits the received sample values to the DSP 63. The correlation discrimination process is defined as a process which judges the similarity between the processed audio data for discrimination which is generated from a series of sample values received from the music CD drive 1 and the processed reference audio data included in the SMF. Hereunder, the correlation discrimination process will be explained with reference to Fig. 8.

[0086]

The DSP 63 receives an execution instruction of the correlation discrimination process from the CPU 62 and records the received sample values in the RAM 64 upon sequentially receiving the sample values, namely, $(r(0), l(0))$, $(r(1), l(1))$, $(r(2), l(2))$, ... Hereunder, a series of 65536 sample values starting from $(r(n), l(n))$ are called as "raw audio data for discrimination (n)". Then, the DSP 63 receives 65536 th sample values, namely, $(r(65535), l(65535))$ and stores the sample value in the RAM 64, and reads out $(r(0), l(0))$ to $(r(65535), l(65535))$, namely, the raw audio data for discrimination (0) from the RAM 64. Subsequently, the DSP 63 executes the correlation discrimination data generation process, which has been mentioned already, namely, the same process as those in the process of step S1 to step S8 in Fig. 4, on the raw audio data for discrimination (0). As a result, the DSP 63 generates 256 sample values and stores the 256 generated sample values in the RAM 64 (step S11). Hereunder, the 256 sample values generated as a result of the correlation discrimination data generation process on the raw audio data for discrimination (n) are represented as $Y_n(0)$ to $Y_n(255)$ and a series of the data is called as "processed audio data for discrimination (n)".

[0087]

Subsequently, the DSP 63 reads out the processed reference audio data included in the system exclusive event in the SMF, namely, X(0) to X(255), and the processed audio data for discrimination (0) stored in the step S11, namely, Y₀(0) to Y₀(255) (step S12) from the RAM 64.

[0088]

Subsequently, the DSP 63 executes a discrimination process represented by following expression (2) and expression (3) (step S13).

[Expression 2]

$$\frac{\sum_{i=0}^{255} (X(i) \times Y_0(i))}{\sum_{i=0}^{255} (X(i)^2)} \geq p \dots\dots (2)$$

[Expression 3]

$$\frac{\left\{ \sum_{i=0}^{255} (X(i) \times Y_0(i)) \right\}^2}{\sum_{i=0}^{255} (X(i)^2) \times \sum_{i=0}^{255} (Y_0(i)^2)} \geq q \dots\dots (3)$$

[0089]

The left side of the expression (2) approaches to 1 as the values of X(m) and Y₀(m) become approximate thereto. Making a pair of data having the identical numbers after sequentially arranging the processed reference audio data and the processed audio data for discrimination (0) in order, the more data values of the respective pairs match, the

larger the left side becomes. In the following explanation, the value of the left side of the expression is called as an absolute correlation index. The value of p is arbitrarily modified within a range of 0 to 1; it is empirically determined in order to acquire an affirmative result (hereinafter referred to as "Yes") when the discrimination is done by the above described expression (2) by using a partial raw reference audio data generated from the same portion of the audio data of the musical tune and the processed audio data for discrimination and in order to acquire a negative result (hereinafter referred to as "No") when the discrimination is done by the expression (2) by using the processed audio data for discrimination acquired from a different portion of the audio data of the musical tune, even though it is similar thereto, and the processed audio data for discrimination.

[0090]

The value of the left side of the expression (3) ranges from 0 to 1 and approaches to 1 as shapes of the audio waveshape represented by $X(m)$ and the waveshape of the audio waveform represented by $Y0(m)$ become more similar. The value of the left side of the expression is called as a relative correlation index in the following explanation. The value of the above mentioned absolute correlation index is smaller than 1 depending on its level, if a level of the audio waveform represented by the processed data for discrimination is lower than a level of the audio waveform represented by the processed reference audio data even though the processed reference audio data and the processed audio data for discrimination are generated from the same portion of the audio data of the musical tune. To the contrary, when the level of the audio waveform represented by the processed audio data for discrimination is large, the absolute relative index becomes larger than 1 depending on the level. On the other hand, since the relative correlation index approximates 1 in any case, the judgment by the expression (3) gives Yes even if the

recording levels are different depending on different versions of the music CDs. The value of q is arbitrarily modified in a range of 0 to 1 and is empirically determined as p .

[0091]

If either or both of results of two judgment in step S13 are No, the DSP 63 finishes the correlation discrimination process using the processed audio data for discrimination (0) and waits for a completion notice of writing a next sample value into the RAM 64 from the CPU 62. The CPU 62 receives a new sample value from the music CD drive 1 (step S14), records it in the RAM 64, and transmits a completion notice of writing the new sample value into the RAM 64 to the DSP 63. The DSP 63 receives the completion notice and the process returns to the above described step S11. However, the data is generated for correlation discrimination for the raw audio data for discrimination having a last sample values of a newly recorded sample values instead of the raw audio data for discrimination (0). As a result, the processed audio data for discrimination (n-1) is recorded in the RAM 64 by a nth time process in the step S11.

[0092]

On the other hand, if both of the results of the two discrimination processes in the step S13 are Yes, the DSP 63 executes discrimination processes represented by following expression (4) and expression (5) (step S15).

[Expression 4]

$$\frac{d \sum_{i=0}^{255} (X(i) \times Y_n(i))}{dn} = 0 \dots (4)$$

[Expression 5]

$$d^2 \sum_{i=0}^{255} (X(i) \times Y_n(i))$$

$$\frac{\Delta^2 n}{d^2 n} < 0 \dots (5)$$

[0093]

The left side of the expression (4) is a variation rate of sum of products of $X(m)$ and $Y_n(m)$ when $n=0$. In the following explanation, the sum of products of $Y(m)$ and $Y_n(m)$ is called as a correlation value. When the processed reference audio data and the processed audio data for discrimination are arranged in order and data having the same order are paired therewith, the more the pair data values become approximate, the larger the correlation value becomes. The variation ratio of the correlation value becomes 0 when the correlation value becomes an extremum after the correlation values, such as the correlation value between $X(m)$ and $Y_0(m)$, the correlation value between $X(m)$ and $Y_1(m)$, ... are arranged in the time axis order. Accordingly, the discrimination process by the expression (4) is a process to judge whether the correlation value is an extremum or not. The process in the expression (5) is to judge whether the extremum is a relative maximum value or not.

[0094]

Since there is no correlation value precedent to a case of $n=0$, the judgment is not enabled. In the present embodiment, the judgment result in the step S15 is No when $n=0$. This is because the raw reference audio data does not start from a head of the music CD-A and the audio data extracted from a timing when the audio waveform represented by the audio data exceeds a threshold and the possibility of the audio data corresponding to the data locating at the head of the music CD-B is extremely low.

[0095]

Explaining more precisely, since $X(m)$ and $Y_n(m)$ are discrete values in the present embodiment, the left side of the expression barely becomes 0. Accordingly, the

judgment process in the step S15 is executed as follows. The DSP 63 makes a difference between a sum of products of $X(m)$ and $Y_n(m)$ and a sum of products of $X(m)$ and $Y_{n-1}(m)$. The value hereunder is called as D_n . Subsequently, the DSP 63 judges whether D_{n-1} is larger than 0 or not D_n is equal to 0 or less. Since when D_{n-1} is larger than 0 and D_n is 0 or less, the variation ratio of the correlation value varies from a positive value to 0 or across 0 at D_n , the correlation value at this time is a relative maximum or an approximate value of the relative maximum. Accordingly, the judgment result in the step S15 is Yes. When the above described process is executed, n needs to be equal to 2 or more and the judgment result in the step S15 is No due to the same reason for the cases that $n=1$ or $n=0$.

[0096]

The judgment result is No in the step S15, the DSP 63 waits for a completion notice of writing a new sample value from the CPU 62. When the CPU 62 receives the completion notice of writing the new sample value (step S14), the DSP 63 returns to the process in the step S11. As a result, a new set of the processed audio data for discrimination in the RAM 64.

[0097]

If the judgment result in the step S13 or the judgment result in the step S15 becomes No to return the process to the step S11 through the Step S14, the DSP 63 continues to process the above mentioned step S12 to step S15. As a result, the DSP 63 sequentially renews the processed audio data for discrimination, such as processed audio data for discrimination (0), processed audio data for discrimination (1), processed audio data for discrimination (2), ... until the judgment result in the step S15 becomes Yes.

[0098]

It is assumed that a musical tune is stored in the music CD-B recorded as the audio data which is delayed from the audio data recorded in the music CD-A with 51,600 samples from the playback start timing, namely, for about 1.17 seconds. In other words, since the audio data (R(52156), L(52156)) to (R(117691), L(117691)) recorded in the music CD-A is extracted as the raw reference audio data, the audio data corresponding to the raw reference audio data in the music CD-B is (r(103756), l(103756)) to (r(169291), l(169291)).

[0099]

In this case, the DSP 63 gets No as a result of judgment in the step S13 or step S15 using the processed audio data for discrimination (0) to the processed audio data for discrimination (103755). This is because the raw audio data for discrimination (0) to the raw audio data for discrimination (103755) used for generating the processed audio data for discrimination do not correspond to the raw reference audio data and the correlation is not enough.

[0100]

The DSP 63 gets Yes as a judgment result in the step S13 executed with the processed audio data for discrimination (103756) and gets Yes as a judgment result in the step S15. This is because the raw audio data for discrimination (103756) used for generating the processed audio data for discrimination (103756) corresponds to the reference raw audio data for enough correlation. As a result, the DSP 63 finishes a series of correlation processes and transmits a notice of success of the correlation discrimination process to the CPU 62.

[0101]

Fig. 9 shows graphs represented by values calculated for samples of actual audio

data with the calculation expressions used in the judgment process in the step S13 and the step S14. Upon creating the graphs, a one-stage IIR (Infinite Impulse Response) filter is used as a high pass filter having a cut-off frequency of 25 Hz in the step S3 in Fig. 4; a combination of $k = 4410$ and $f = 1$ are used as constants in the comb filter in the step S5; and a one-stage IIR filter is used as a low pass filter having a cut-off frequency of 25 Hz in the step S6. Further, constants of $p = 0.5$ and $q = 0.8$ are used in a criterion in the step S13.

[0102]

The graph at the top of Fig. 9 shows the values of the numerator of the left side of the expression (2) and values in the expression which the denominator in the left side is moved to the right side with respect to n (abscissa). The middle graph in Fig. 9 shows values of the numerator of the left side of the expression (3) and values in the expression which the denominator in the left side is moved to the right side with respect to n . The bottom graph in Fig. 9 shows the values in the left side in the expression (4).

[0103]

When the value of n is within a domain A in Fig. 9, the value of the numerator at the left side of the expression (2) is equal to or greater than the value of the expression which the denominator of the left side is moved to the right side and the condition of the Expression (2) is met. In the domain A, when n is located in a domain B, the value of the numerator of the left side of the Expression (3) is equal to or greater than the value of the expression which the denominator of the left side of the Expression (3) is moved to the right side and the condition of the Expression (3) is met. As a result, the affirmative result (Yes) is got in the judgment process in the step S13. When the value of n is equal to a value as indicated with an arrow C in the domain B, the value of the left side of the

Expression (4) turns from a positive value to 0 and the condition of the Expression (5) is met; an affirmative result (Yes) is got in the judgment process in the step S15.

[0104]

[1.2.2.4] Playback of MIDI event

The CPU 62 receives a success notice of the correlation discrimination process from the DSP 63 and starts time measurement by setting that timing as 0 second. At the same time, the CPU 62 reads out SMF from the RAM 64 and sequentially compares the measured time with the delta time included in the SMF; when the measured time coincides with the delta time, the MIDI event corresponding the delta time is transmitted to the automatic player piano 3.

[0105]

In the automatic player piano 3, the MIDI event control circuit 34 receives the MIDI events from the CPU 62 and transmits the received MIDI events to the tone generator 35 or the driving part 36. When the MIDI events are transmitted to the tone generator 35, the tone generator 35 sequentially transmits audio data indicative of tones of a musical instrument based on the received MIDI events to the tone generating portion 4. The tone generating portion 4 outputs the sounds of the musical tune in the music CD-B which is played back already and the performance by the musical tone received from the tone generator 35 from the speaker 44. On the other hand, when the MIDI events are transmitted to the driving portion 36, the driving portion 36 drives the keys and pedals of the piano 31 based on the received MIDI events. In either case, the user simultaneously listens to the musical tune recorded on the music CD-B and the performance with the musical tone by the performance information recorded in the SMF.

[0106]

[1.2.2.5] Time relation between audio data and MIDI event

As described in the above, the user simultaneously played back the music CD and the MIDI events recorded in the SMF, however, the time gap between the starting timings of the musical tune in the music CD-A and the music CD-B are adjusted and the music CD and the MIDI events recorded in the SMF are simultaneously played back. Time relation between the audio data in the music CD-A, the music CD-B and the MIDI events are summarized in Fig. 10. In Fig. 10, it is illustrated that the level of the audio waveform represented by the audio data of the music CD-B is generally lower than the audio waveform represented by the audio data of the music CD-A. To discriminate two different timings, the time starting at a playback starting time of the music CD-B which is set as 0 second is labeled with (T') as a suffix.

[0107]

When the gap between the starting timings of the musical tune in the music CD-A and the music CD-B are not adjusted and the MIDI events are played back based on the playing start timing of the music CD, a first event, a second event and a third event are transmitted to the automatic player piano 3 at 3.92 seconds (T'), 5.30 seconds (T') and 6.38 seconds (T'), respectively. Accordingly, the performance by the MIDI events runs earlier than the musical tune in the music CD.

[0108]

Since the raw audio data for discrimination extracted from the music CD-B and the raw reference audio data already extracted from the music CD-A are different very much for about 3.84 seconds after the playback of the music CD-B is started, there is insufficient correlation between the processed audio data for discrimination and the processed reference audio data generated therefor, respectively, so that the playback of

MIDI events is not started.

[0109]

The correlation is sufficient between the sets of audio data at about 3.84 seconds (T') and it is judged that each set is generated from the same portion of the musical tune in the music CD-B and the music CD-A. The measurement of the delta time is started at about 3.84 seconds (T'), namely, about 2.67 seconds (T) in the music CD-A, and a first event, a second event and a third event are transferred to the automatic player piano 3 at about 5.09 seconds (T'), 6.47 seconds (T') and 7.55 seconds (T'), respectively. In such a way, the transmission timings of the MIDI events are adjusted and the performance based on the MIDI events is generated for the musical tune recorded in the music CD-B.

[0110]

[2] Second embodiment

In a second embodiment of the present invention, time codes recorded in a music CD are used for synchronized playback of the audio data recorded in the music CD and MIDI events recorded in the SMF.

[0111]

[2.1] Music CD drive

Since the whole configuration, the functions of respective elements and the data format of the MIDI data of the second embodiment are similar to those in the first embodiment except a function of the music CD drive 1, the function of the music CD drive 1 is explained only and other explanation will be omitted.

In the second embodiment, the music CD drive 1 transmits the time codes to the controller 6 together with the audio data recorded in the music CD. Other features are the same as those of the music CD drive 1 in the first embodiment.

[0112]

[2. 2] Operation

The operation of the synchronized recorder and player SS in the second embodiment differs from the first embodiment in following three points.

- (1) A time code at a starting timing of the raw reference audio data used for generating the processed reference audio data is recorded in a system exclusive event in the SMF.
- (2) A time code corresponding to a generation timing of the MIDI events is recorded as a delta time of another MIDI event recorded in the SMF.
- (3) The MIDI event is not measured based on a clock signal by the controller 6 during the playback of the MIDI events and is transmitted to the automatic player piano 3 based on the time clock transmitted from the music CD drive 1.

[0113]

Other operations in the second embodiment are similar to those of the first embodiment and the detailed explanation will be omitted. It is assumed that the music CD-A and the music CD-B are used for a recording operation and a playback operation, respectively, as similar to the first embodiment in the following explanation. The format of the time code is represented by hour, minute, second and frame and the time information represented by the time code is represented in second as similar to the delta time recorded in the SMF in the following explanation for the sake of the simplicity.

[0114]

[2.2.1] Recording operation

In the synchronized recorder and player SS in the second embodiment, when the user instructs the recording start of the performance data through the manipulation display 5, the audio data in the music CD-A is sequentially transmitted to the controller 6 from the

music CD drive 1 together with the time codes.

[0115]

In the controller 6, the CPU 62 sequentially transmits the received audio data to the tone generating portion 4 and the musical tune of the music CD-A is output as sounds from the tone generating portion 4. On the other hand, if the absolute value of the sample value of the received audio data exceeds a predetermined threshold, the CPU 62 converts the time code which is received immediately before, into the format of the delta time and stores the data in the RAM 64. In other word, the RAM 64 stores "1.18 seconds" as the delta time. Hereunder, the delta time is called as a "reference audio data starting time".

[0116]

The CPU 62 records the reference audio data starting time and, at the same time, starts recording the sample values in the RAM 64, and after that, sample values for about 1.49 second are stored in the RAM 64 as the raw reference audio data.

After finishing recording of the raw reference audio data by the CPU 62, the DSP 63 reads out the raw reference audio data from the RAM 64 and executes the data generation process for correlation discrimination on the readout raw reference audio data. As a result, the processed reference audio data is stored in the RAM 64.

[0117]

The DSP 63 executes the data generation process for correlation discrimination and the user starts playing the piano 31 with sounds of the musical tune in the music CD-A listened through the tone generating portion 4. The performance information by the user is transmitted from the automatic player piano 3 to the controller 6 as the MIDI events. The CPU 62 receives the MIDI events and converts the time code, which is received from the music CD drive 1 immediately before, into the format of the delta time and stores the

data into the RAM 64 in association with the MIDI event.

[0118]

After the musical tune of the music CD-A is finished and the performance by the user is also finished, the user instructs to finish the recording the performance data through using the manipulation display 5. After the instruction by the user, the playing of the music CD-A by the music CD drive 1 is stopped. Subsequently, the CPU 62 reads out the reference audio data starting time, the processed reference audio data, the MIDI event generated through the performance by the user and the delta time associated with the MIDI event from the RAM 64. The CPU 62 combines these data which are read out in order to generate the SMF.

[0119]

Fig. 11 is a view to show the overview of the SMF generated by the CPU 62. The SMF stores the processed reference audio data and the reference audio data starting time in the system exclusive event. The delta time corresponding to another MIDI event includes the same time information as the time code substantially simultaneously received by the CPU 62 and the delta time for the first event is 3.92 seconds, for example. The delta time indicates that the first event is generated at 3.92 seconds after the playing start of the audio data of the music CD-A.

The CPU 62 transmits the generated SMF together with a write instruction to the FD drive 2 and the FD drive 2 writes the SMF in the FD.

[0120]

[2.2.2] Playback operation

Subsequently, operations for playing back the SMF, which is recorded by the above mentioned method, and for synchronizing the audio data of the music CD-B and the

MIDI data of the SMF will be explained.

The user instructs the playback start of the performance data by using the manipulation display 5; the SMF recorded in the FD is transmitted from the FD drive 2 to the CPU 62 and the CPU 62 stores the received SMF in the RAM 64. Subsequently, the music CD drive 1 starts playback of the music CD-B and the audio data and the time code recorded in the music CD-B are sequentially transmitted to the controller 6. The CPU 62 sequentially transmits the received audio data to the tone generating portion 4 and the musical tune of the music CD-B is output from the tone generating portion 4 as sounds. The CPU 62 sequentially transmits the audio data to the tone generating portion 4 and at the same time, records the audio data together with the time code in the RAM 64.

[0121]

When the CPU 62 records a 65536th sample value in the RAM 64, the DSP 63 starts the correlation discrimination process for the audio data recorded in the RAM 64. Then, the DSP 63 generates the process audio data for discrimination from the raw audio data for discrimination, which is sequentially renewed, and repeats the judgment processes in the step S13 and the step S15 on the processed audio data for discrimination already generated until the judgment result becomes Yes in the step S15 in Fig. 8.

[0122]

The DSP 63 gets Yes as the judgment result in the step S15 by using the processed audio data for discrimination (103756), and a series of correlation discrimination process is finished and a success notice of the correlation discrimination process is transmitted to the CPU 62. The success notice of the correlation discrimination process includes a number "103756" of the processed audio data for discrimination (103756) finally used in the correlation discrimination result.

[0123]

When the CPU 62 receives the success notice of the correlation discrimination result from the DSP 63 and reads out the head sample value of the raw audio data for discrimination (103756), namely, (r(103756), l(103756)) together with the time code stored in the RAM 64. In this case, the time indicated by the time code is 2.35 seconds. Subsequently, the CPU 62 calculates the difference between the time represented by the readout time code and the reference audio data starting time included in the system exclusive event in the SMF stored in the RAM 64.

[0124]

In this case, the time represented by the reference audio data starting time is 1.18 seconds, and the time difference is 1.17 seconds. It shows that the delta time recorded in the SMF is 1.17 seconds earlier than the musical tune of the music CD-B as a whole. Accordingly, the CPU 62 adds 1.17 seconds to respective delta time in the SMF. As a result, the delta time for the first event, the second event and the third event are renewed from 3.92 seconds, 5.30 seconds and 6.38 seconds to 5.09 seconds, 6.47 seconds and 7.55 seconds, respectively. Hereunder, this operation is called as a "timing adjustment process".

[0125]

Subsequently, the CPU 62 sequentially compares the time code of the music CD-B, which is sequentially transmitted from the music CD drive 1, with the renewed delta time, and the MIDI event corresponding to the delta time is transmitted to the automatic player piano 3 when the pair of the time information matches each other.

[0126]

The automatic player piano 3 executes the automatic performance based on the

MIDI events transmitted from the controller 6. As a result, the user simultaneously listens to the musical tune recorded in the music CD-B and the performance based on the performance information recorded in the SMF.

[0127]

[2.2.3] Time relation between audio data and MIDI event

Fig. 12 is a view to show a relation with respect to time between the audio data in the music CD-A and the music CD-B, and the MIDI data during the recording operation and the playback operation of the MIDI data.

The upper view in Fig. 12 shows a relation between the time represented by the time code in the music CD-A during the recording operation for the MIDI data and the time represented by the delta time associated with the recorded MIDI data. As shown in the drawing, the time information represented by the time code upon the generation of the MIDI event is recorded in the delta time as it is.

[0128]

The middle view in Fig. 12 shows a relation between the time represented by the time code in the music CD-B during the playback operation of the MIDI data and the time represented by the delta time after the timing adjustment process. When the MIDI events are played back by using the delta time before the timing adjustment process based on the time code in the music CD-B, the MIDI events are played back earlier than the musical tune in the music CD-B. However, the time difference is adjusted through the timing adjustment process; the MIDI events are played back by using the delta time after the timing adjustment process based on the time code in the music CD-B; and the MIDI events are played at correct timings with respect to the musical tune in the music CD-B.

[0129]

Incidentally, the music CD drive 1 divides a basic clock signal from an oscillator included in the music CD drive 1 in order to generate a clock signal at 44100 Hz and sequentially transmits the audio data recorded in the music CD to the controller 6 based on the clock signal. If the operation of the oscillator is unstable, the playback speed is slightly different every time when the same music CD is played.

[0130]

The lowest view in Fig. 12 shows a relation between the time represented by the time code in the CD-B and the time represented by the delta time after timing adjustment process when the music CD-B is played back at a playback speed slightly higher than the playback speed for playing the music CD-B in the middle view. If the MIDI event is reproduced based on the clock signal of the CPU 62, the playback of the MIDI events are slightly delayed from the music CD-B. In other words, it is assumed that the middle view in Fig. 12 is based on the time according to the clock signal of the CPU 62 and the clock signal of the CPU 62 and the division process is free of error, a first event, a second event and a third event are played later than the music CD-B for t1, t2 and t3, respectively, due to the errors in the clock signal and the division process in the music CD drive 1.

[0131]

Since the MIDI events are played back based on the time codes transmitted from the music CD drive 1 to the CPU 62 in real time in the second embodiment, the MIDI event is not played back by being delayed from the musical tune in the music CD-B with the time difference.

[0132]

[3] Third embodiment

In a third embodiment of the present invention, the raw reference audio data is

extracted from the middle of the musical tune instead of the start of the musical tune represented by the audio data recorded in a music CD. Further, in the third embodiment, a time code recorded in the music CD is used to synchronously adjust playing MIDI events recorded in an SMF as similar to the second embodiment.

Since the whole configuration, functions of respective elements and the data format of the MIDI data in the third embodiment are similar to those in the second embodiment, the explanation on those will be omitted.

[0133]

[3.1] Operation

The operation of the synchronized recorder and player SS of the third embodiment is different from the second embodiment in following two points.

(1) The raw reference audio data is extracted from the middle of the musical tune represented by the audio data recorded in the music CD during the recording operation of the MIDI event.

(2) During the playback operation of the MIDI event, a playback timing of the MIDI event is determined by the correlation discrimination process on the audio data recorded in the music CD, and after that, the audio data recorded in the music CD is played back from a start.

[0134]

[3.1.1] Recording operation

In the recording operation of the MIDI event in the third embodiment, an arbitrary portion of the audio data recorded in the music CD is extracted as the raw reference audio data. For example, sample values for about 1.49 seconds from a timing after 3 minutes lapsing after the start of the musical tune may be the raw reference audio data or sample

values for about 1.49 seconds including a featured audio waveform in the whole musical tune may be the raw reference audio data. In the following explanation, the raw reference audio data for about 1.49 seconds from a timing after 3 minutes, namely, 180 seconds in the time code of the musical tune in the music CD-A, as an example.

[0135]

If the user instructs the recording start of the performance data, 65536 pairs of audio data for 180 seconds from the starting point of the musical tune of the music CD-A are transmitted to the CPU 62 from the music CD drive. The CPU 62 converts the time code at the head of the received audio data into the delta time format and records the data in the RAM 64 as the reference audio data starting time. The CPU 62 records the sample values of the audio data included in the received audio data in the RAM 64 as the raw reference audio data. The CPU 62 executes the correlation discrimination data generation process on the raw reference audio data and records the processed reference audio data in the RAM 64 as a result of this.

[0136]

Subsequently, the music CD drive 1 plays back the music CD-A from a start. The CPU 62 sequentially receives the audio data from the music CD drive 1 and transmits the audio data included in the received audio data to the tone generating portion 4. The user performs with the piano 31 in ensemble with sounds of the musical tune of the music CD-A generated from the tone generating portion 4, and the performance information is sequentially transmitted to the CPU 62 as the MIDI events. The CPU 62 receives the MIDI events in order to convert the time code received from the music CD drive 1 immediately before into the delta time format and stores the data in the RAM 64 in association with the MIDI event.

[0137]

The user instructs the recording finish of the performance data and the music CD drive 1 stops playing the music CD-A. At the same time, the CPU 62 creates an SMF shown in Fig. 13 from the data stored in the RAM 64. The created SMF is written in an FD by the FD drive 2.

[0138]

[3.1.2] Playback operation

Subsequently, when the SMF recorded by the above mentioned method is played back in synchronization with the music CD-B, the SMF is transmitted to the CPU 62 from the FD drive 2 by the instruction of playing start of the performance data by the user. The SMF is stored in the RAM 64. Subsequently, the audio data and the time codes of the CD-B are sequentially transmitted from the music CD drive 1 to the CPU 62.

[0139]

The CPU 62 receives the sample value of the 65536th audio data and starts the correlation discrimination process on the received series of sample values. Under the control of the CPU 62, the DSP 63 renews the raw reference audio data to be used for the correlation discrimination process with the sample values of the audio data, which are sequentially received, and repeats the correlation discrimination process until the judgment result of the step S15 in Fig. 8 becomes Yes. Since the musical tune of the music CD-B is delayed from the musical tune of the music CD-A for about 1.17 seconds, the CPU 62 receives the sample value corresponding to a timing about 182.35 seconds after the head of the musical tune of the music CD-B and executes the correlation discrimination process on the raw reference audio data having this sample value as the final one, and the judgment result of the step S15 results in Yes to finish the correlation discrimination result. The

CPU 62 acquires 181.17 seconds as a time code corresponding to the head data of the raw reference audio data used upon succeeding the correlation discrimination process by the CPU 62.

[0140]

The CPU 62 calculates a difference between the time represented by the time code and the time represented by the delta time included in the system exclusive event of the SMF. In this case, the time difference between the time is 1.17 seconds and the CPU 62 adds 1.17 seconds to the respective delta time in the SMF. As a result, the respective delta time is adjusted in order to have a correct timing with respect to the musical tune of the music CD-B as similar to the second embodiment. The above is a process to determine a playback timing of the MIDI events; the musical tune of the music CD-B is not transmitted to the tone generating portion 4 during the process, and accordingly, the musical tune of the music CD-B is not listened by the user.

[0141]

After finishing the above process, the music CD drive 1 reproduces the music CD-B from a start of the musical tune again. The audio data of the musical tune in the music CD-B is transmitted to the tone generating portion 4 through the CPU 62 and the user listens to the musical tune from the tone generating portion 4. At the same time, the CPU 62 sequentially compares the time code of the music CD-B received from the music CD drive 1 with the renewed delta time in the SMF, and if both of the time information coincides therewith, the MIDI event corresponding to the delta time is transmitted to the automatic player piano 3. As a result, the automatic performance is played by the automatic player piano 3.

[0142]

Fig. 14 is an illustrative view of relations between the raw reference audio data, the processed reference audio data, the raw audio data for discrimination and the processed audio data for discrimination in the third embodiment. The raw reference audio data is created by extracting the audio data for about 1.49 seconds from a timing passing a time period T1 from the head of the musical tune CD-A. The raw reference audio data experiences the data generation process for correlation discrimination to create the processed reference audio data. The processed reference audio data is stored in the head of the SMF with the time information representing the time period T1.

The audio data corresponding to the raw reference audio data in the music CD-A is stored as the audio data for about 1.49 seconds from a timing passing a time period T2 from the head in the music CD-B.

[0143]

The adjustment of the delta time of the MIDI event included in the SMF is executed by a difference between T1 and T2. In other words, if T1 is smaller than T2, the difference is added to the delta time in the SMF, and if T1 is larger than T2, the difference is subtracted from the delta time in the SMF.

[0144]

[4] Modifications

The above mentioned first embodiment, second embodiment and third embodiment are mere illustrations of the embodiments of the present invention, and various modifications are available without departing from the feature of the present invention. Modifications will be shown hereunder.

[0145]

[4. 1] First modification

In the first modification, elements of the synchronized recorder and player SS are not located in the same device and are separated into groups to be located.

For example, they are separable into following respective groups:

- (1) music CD drive 1
- (2) FD drive 2
- (3) automatic player piano 3
- (4) mixer 41 and D/A converter 42
- (5) amplifier 43
- (6) speaker 44
- (7) manipulation display 5 and controller 6

Further, the controller 6 may be separated into a device for recording operations only and a device for playback operations only.

[0146]

The element groups are connected with audio cables, MIDI cables, optical audio cables, USB (Universal Serial Bus) cables and dedicated control cables. The FD drive 2, the amplifier 43 and speakers 44 may be commercially available ones.

According to the first embodiment, the location flexibility of the synchronized recorder and player SS is enhanced and the user does not need to prepare the whole new components of the synchronized recorder and player SS to reduce the cost.

[0147]

[4.2] Second modification

In a second modification, the synchronized recorder and player SS does not include the music CD drive 1 and the FD drive 2. On the other hand, a communication interface has a function connectable to the LAN (Local Area Network) and is connected to

external communication devices through the LAN and WAN. The controller 6 has an HD (Hard Disk).

[0148]

The controller 6 receives the digital audio data including the audio data and the time codes from other communication devices through the LAN and records the received audio data in the HD. As similarly, the controller 6 receives the SMF created in association with the audio data from other communication devices through the LAN and records the received SMF in the HD.

[0149]

The controller 6 reads out the digital audio data from the HD instead of receiving the audio data and the time codes of the music CD from the music CD drive 1. The controller 6 executes the similar operations on the HD instead of writing and reading out the SMF into or from the FD drive 2.

According to the second modification, the user is capable of transmitting and receiving the digital audio data and the SMF through the LAN to the communication device which is geographically remote therefrom. The LAN may be connected to the wide area communication network such as the Internet.

[0150]

[4.3] Third modification

In the above mentioned embodiments, all of the discrimination by the absolute correlation index, the discrimination by the relative correlation index and the discrimination by the correlation values are used in the step S14 and the step S15 of the correlation discrimination process, however, the correlation discrimination process is executed by one of or plural combinations of these discriminations in the third embodiment.

One of or plural combinations of these discriminations may be freely selectable.

According to the third modification, the discrimination result having the necessary accuracy is acquired with more flexibility.

[0151]

[4.4] Fourth modification

Though the relative maximum of the correlation value is detected by the discriminations expressed by the expression (4) and the expression (5) in the step S15 in the correlation discrimination process in the above mentioned embodiments, the discrimination expressed by the expression (4) is executed only and the extremum of the correlation value is detected.

[0152]

More specifically, the DSP 63 calculates the product of D_{n-1} and D_n and judges if the product is 0 or less. If the product is 0 or less, the variation ratio of the correlation value is 0 or varies across 0 and the correlation value at this time is an extremum or an approximate value of the extremum. Accordingly, if the product of D_{n-1} and D_n is 0 or less the judgment result in the step S15 becomes Yes.

[0153]

According to the fourth modification, if it is less possible to have a relative minimum value near the relative maximum value, the discrimination result similar to the step S15 in the above mentioned embodiment is acquired by a simpler judgment process.

[0154]

[EFFECTS OF THE INVENTION]

As explained in the above, according to the present invention, the synchronized play back of the performance data is allowed at a correct timing for the audio data of

difference versions in which the audio data having different starting points of the same musical tune. Accordingly, a different set of performance data is not necessary for a different version of the same musical tune and the data generation and data management are simplified.

[0155]

A different version of the same musical tune may have a different recording level of the musical tune; the present invention uses an index representing the similarity between a shape of the audio waveform representing the reference audio data and a shape of the audio waveform representing the actual audio data as an index used for determining the playback start timing of the performance data; the correct playing start timing is determined for the audio data for versions having different recording levels.

[0156]

In the present invention, when the performance data is played back based on the time codes, the performance data is played back at a correct timing with respect to the audio data even when the playback speed of the audio data is unstable.

[BRIEF DESCRIPTION ON DRAWINGS]

[Fig. 1] The block diagram showing the configuration of the synchronized recorder and player SS implemented by the first embodiment and the second embodiment of the present invention.

[Fig. 2] The view to show the data format of MIDI event.

[Fig. 3] The view to show the data format of SMF.

[Fig. 4] The flowchart of the correlation data generation process for discrimination implemented by the first embodiment and the second embodiment of the present invention.

[Fig. 5] The view to show the configuration of the comb filter implemented by the first embodiment and the second embodiment of the present invention.

[Fig. 6] The view to show the relation between the audio data and the MIDI events with respect to time during the recording operation implemented by the first embodiment of the present invention.

[Fig. 7] The view to show the overview of the SMF implemented by the first embodiment of the present invention.

[Fig. 8] The flowchart of the correlation discrimination process implemented by the first embodiment and the second embodiment of the present invention.

[Fig. 9] The view to show the relation between the variation of values of calculation expressions and the discrimination result implemented by the first embodiment and the second embodiment of the present invention.

[Fig. 10] The view to show the relation between the audio data during the recording operation, the audio data during playback operation and MIDI events with respect to time implemented by the first embodiment of the present invention.

[Fig. 11] The view to show overview of the SMF implemented by the second embodiment of the present invention.

[Fig. 12] The view to show the relation between the audio data during the recording operation, the audio data during the playback operation and MIDI events with respect to time implemented by the second embodiment of the present invention.

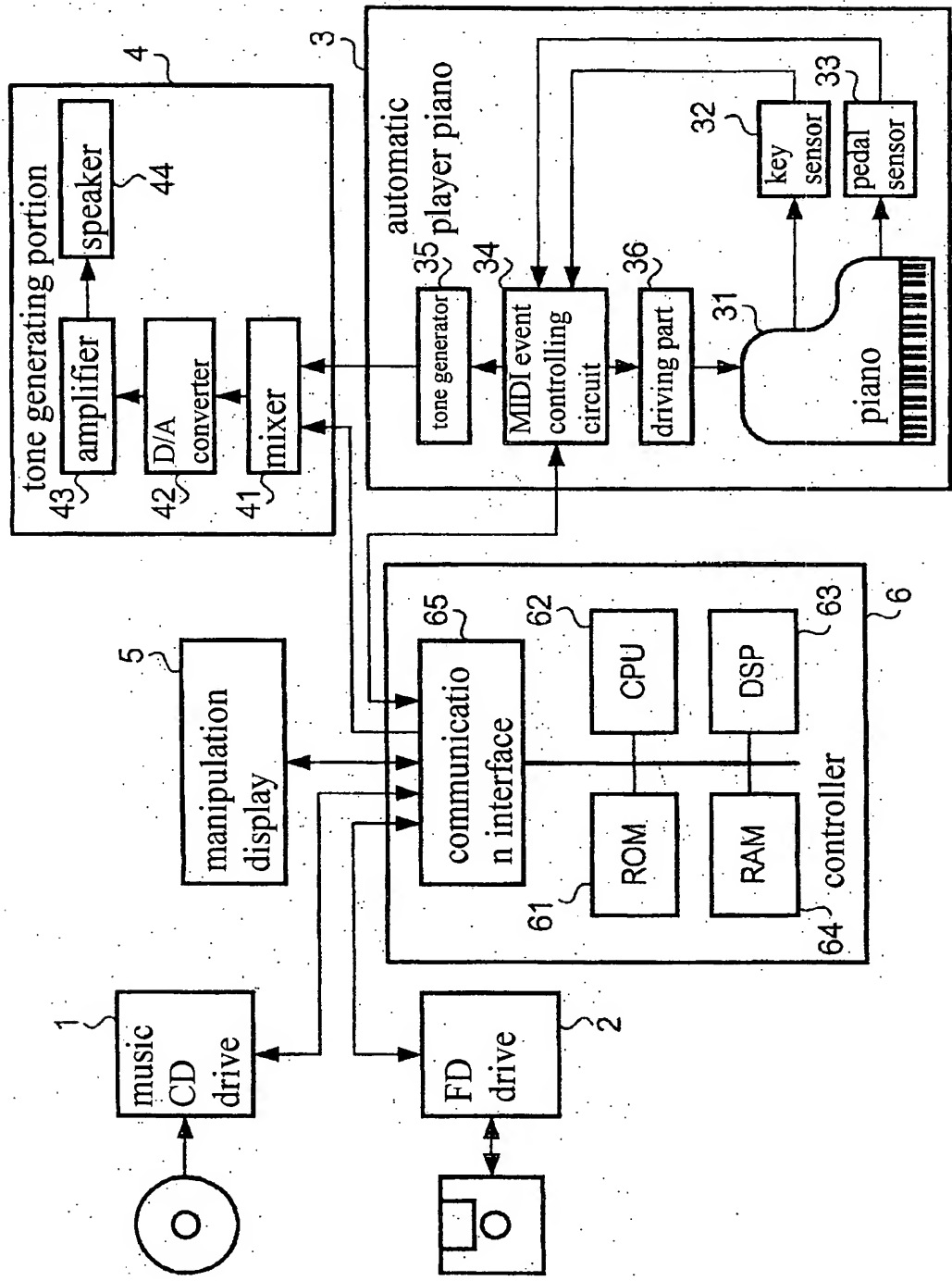
[Fig. 13] The view to show the overview of the SMF implemented by the third embodiment of the present invention.

[Fig. 14] The view to show the raw reference audio data, the processed reference audio data, the raw audio data for discrimination, the processed audio data for

discrimination implemented by the third embodiment of the present invention.

[EXPLANATION ON REFERENCES]

1 ... music CD drive, 2 ... FD drive, 3 ... automatic player piano, 4 ... tone
generating portion, 5 ... manipulation display, 6 ... controller, 31 ... piano, 32 ...
key sensor, 33 ... pedal sensor, 34 ... MIDI event control circuit, 35 ... tone generator,
36 ... driving portion, 41 ... mixer, 42 ... D/A converter, 43 ... amplifier, 44 ...
speaker, 61 ... ROM, 62 ... CPU, 63 ... DSP, 64 ... RAM, 65 ... communication
interface



Note-on event

Note-on (9nH)	note number	velocity
------------------	-------------	----------

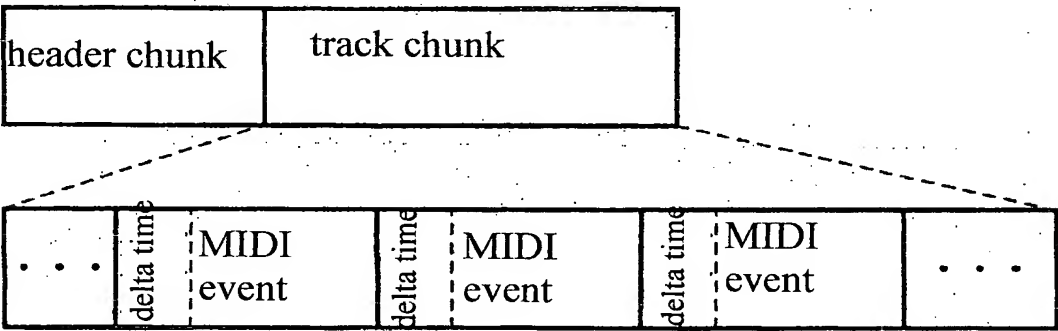
Note-off event

Note-off (8nH)	note number	velocity
-------------------	-------------	----------

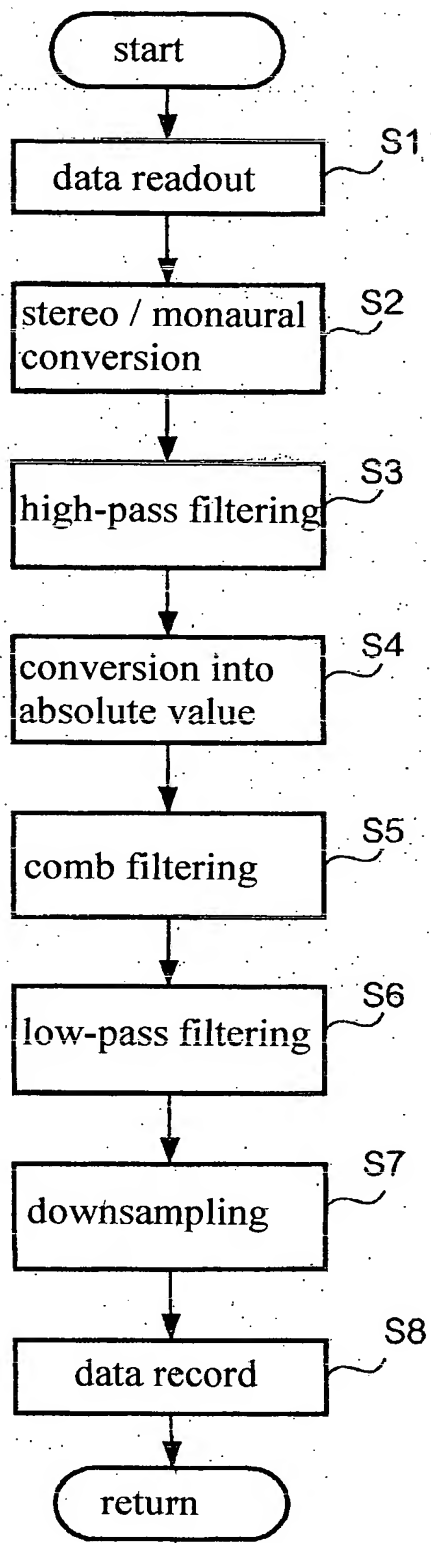
system exclusive event

system exclusive event (F0h)	data length	data	system exclusive event (F7h)
------------------------------------	-------------	------	------------------------------------

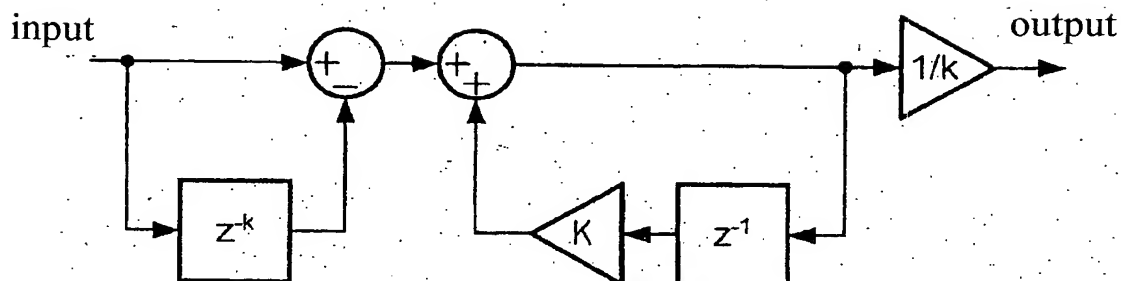
[Fig. 3]



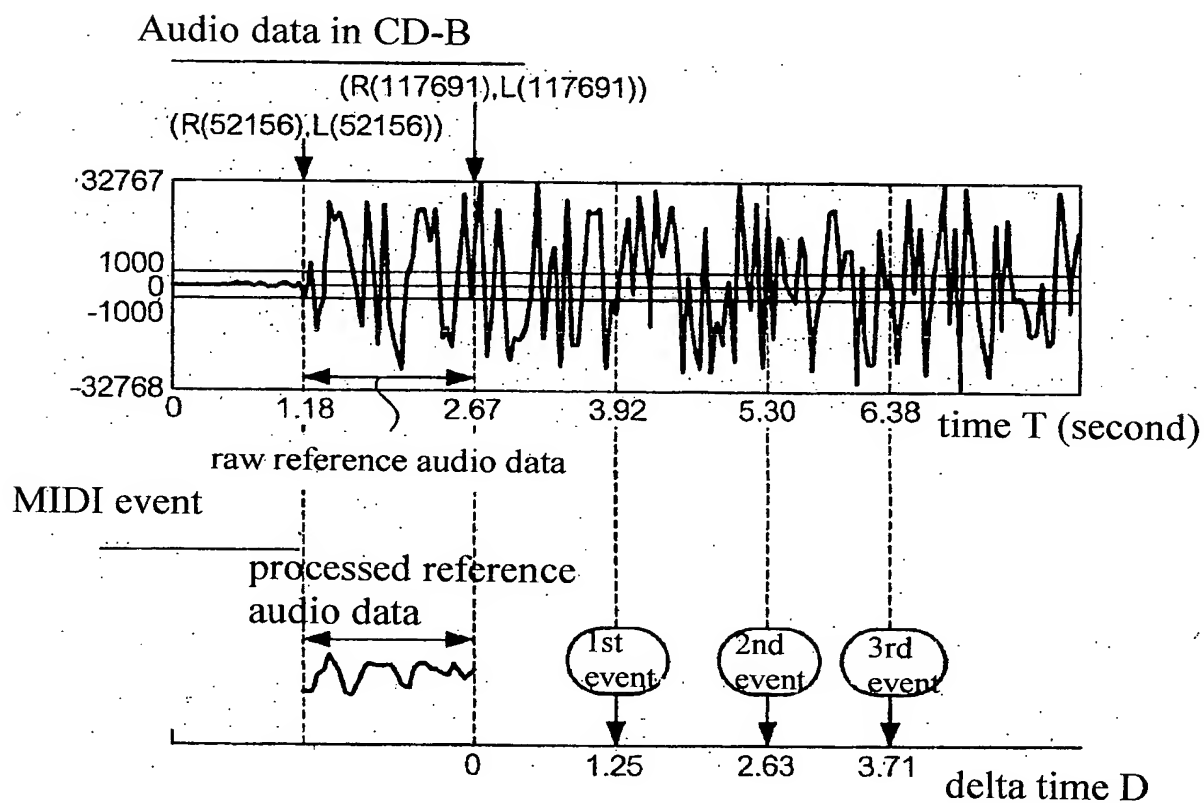
[Fig. 4]



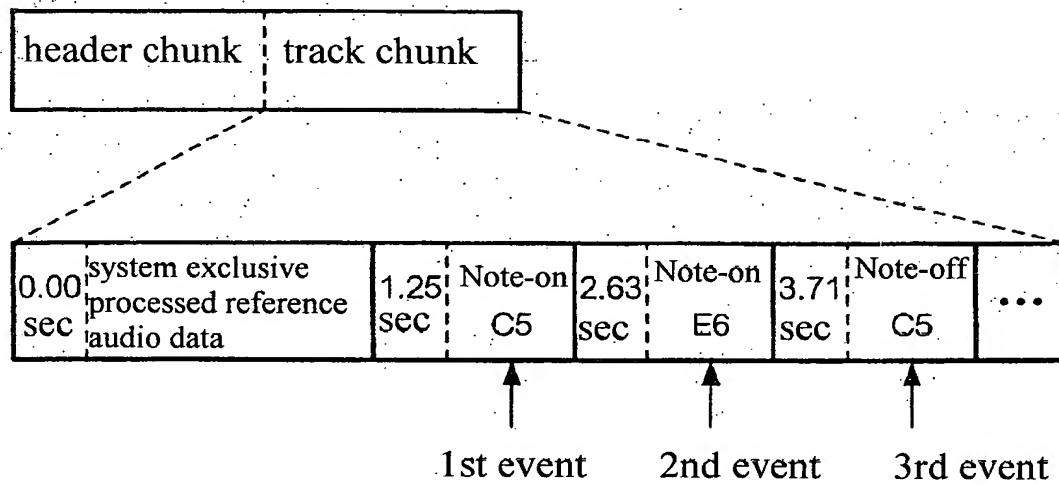
[Fig. 5]



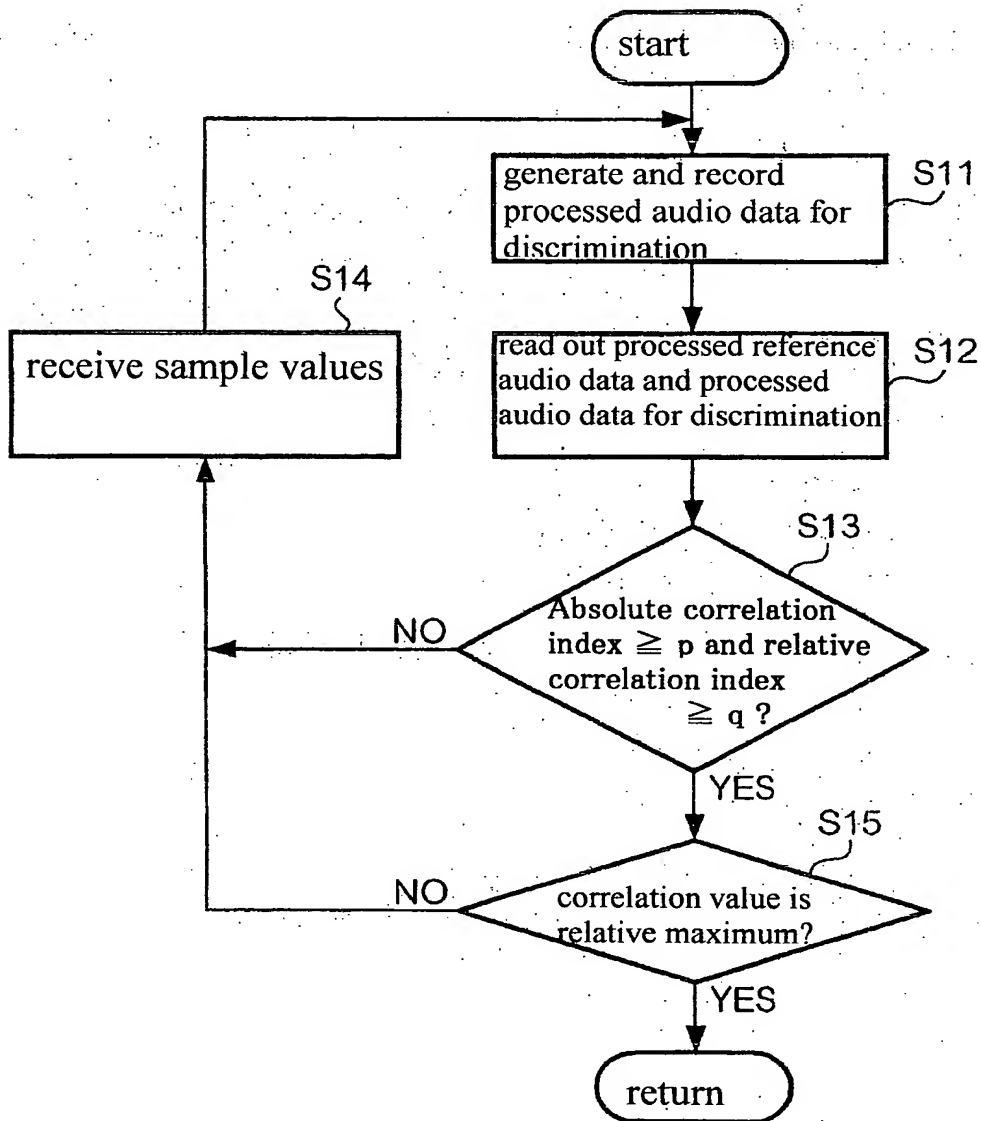
[Fig. 6]

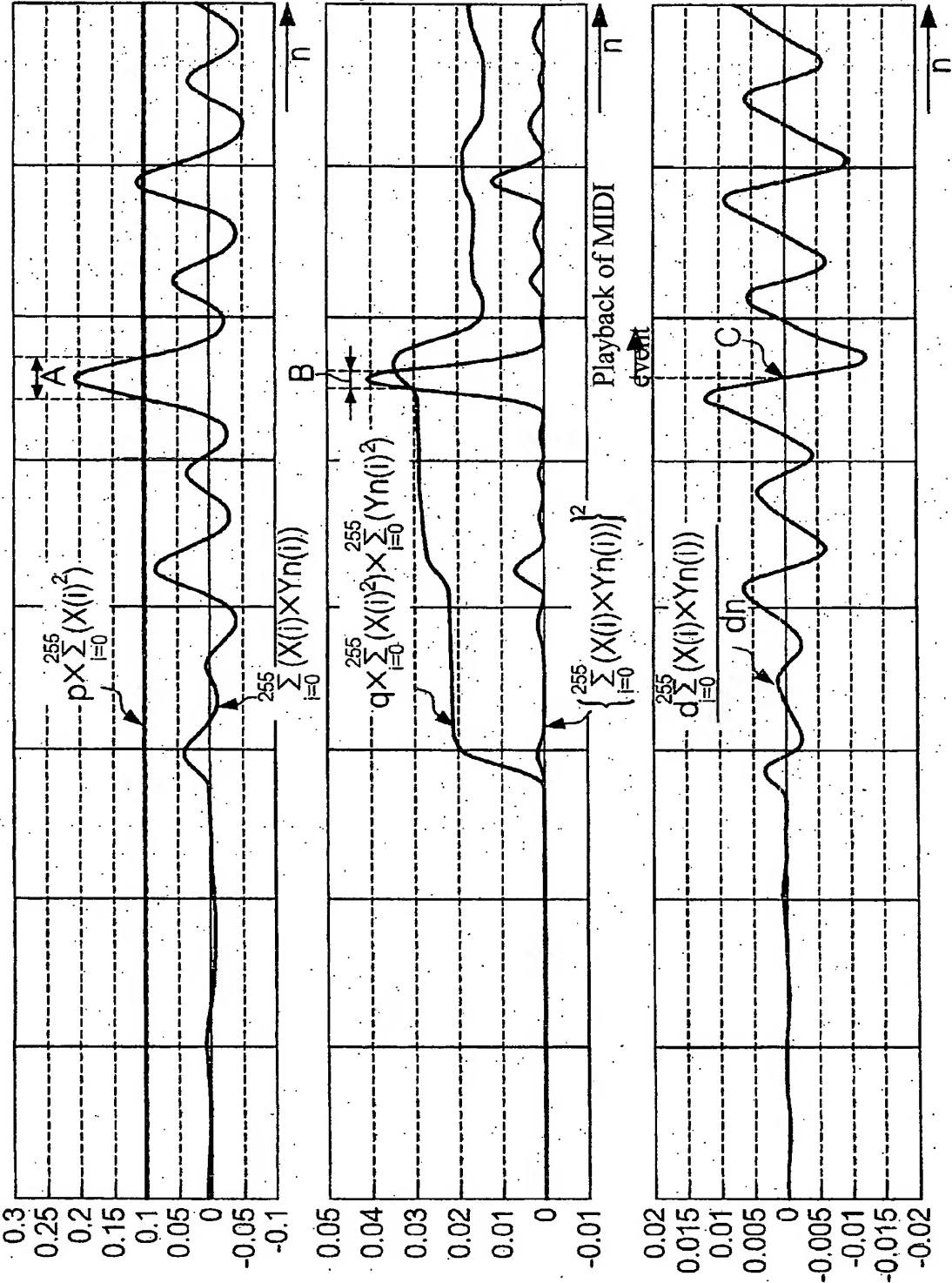


[Fig. 7]

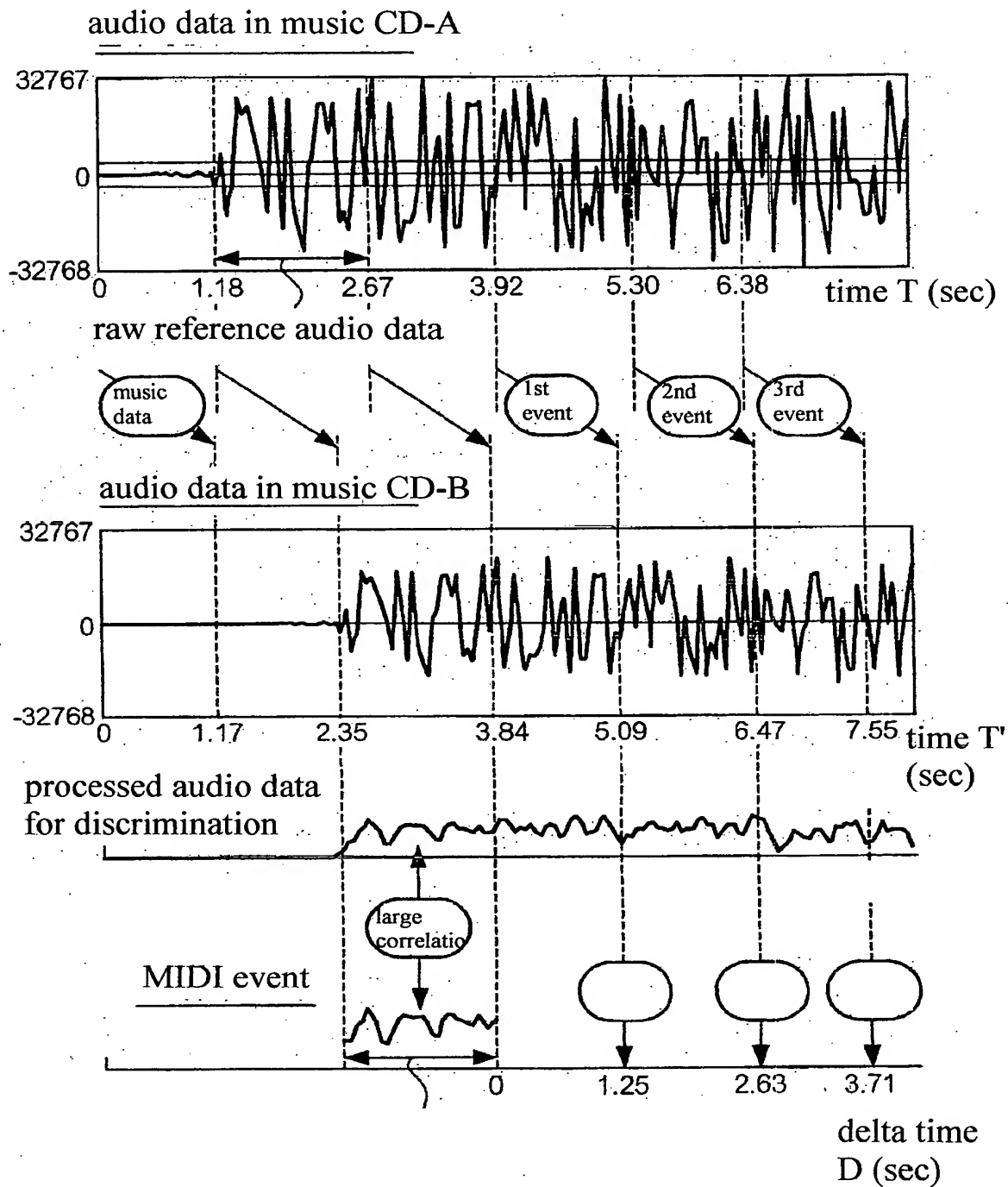


[Fig. 8]

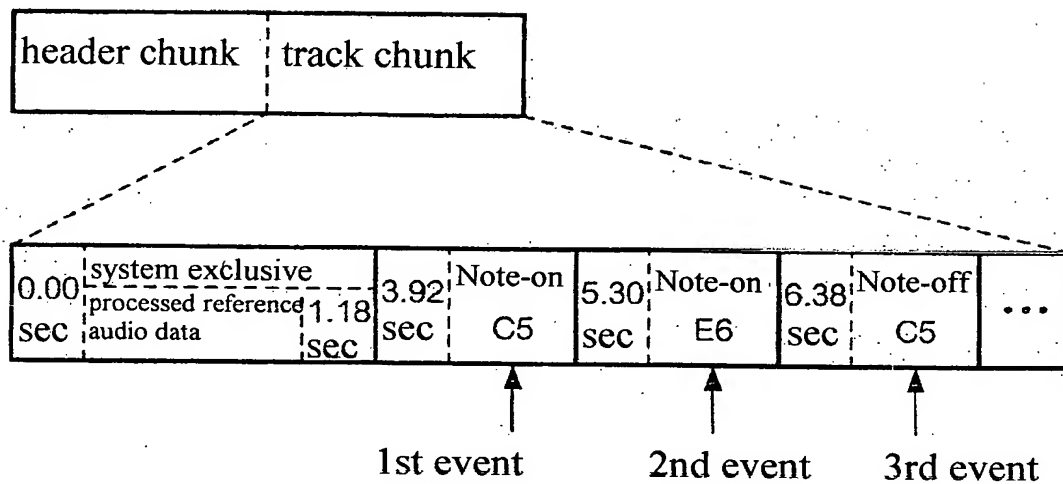




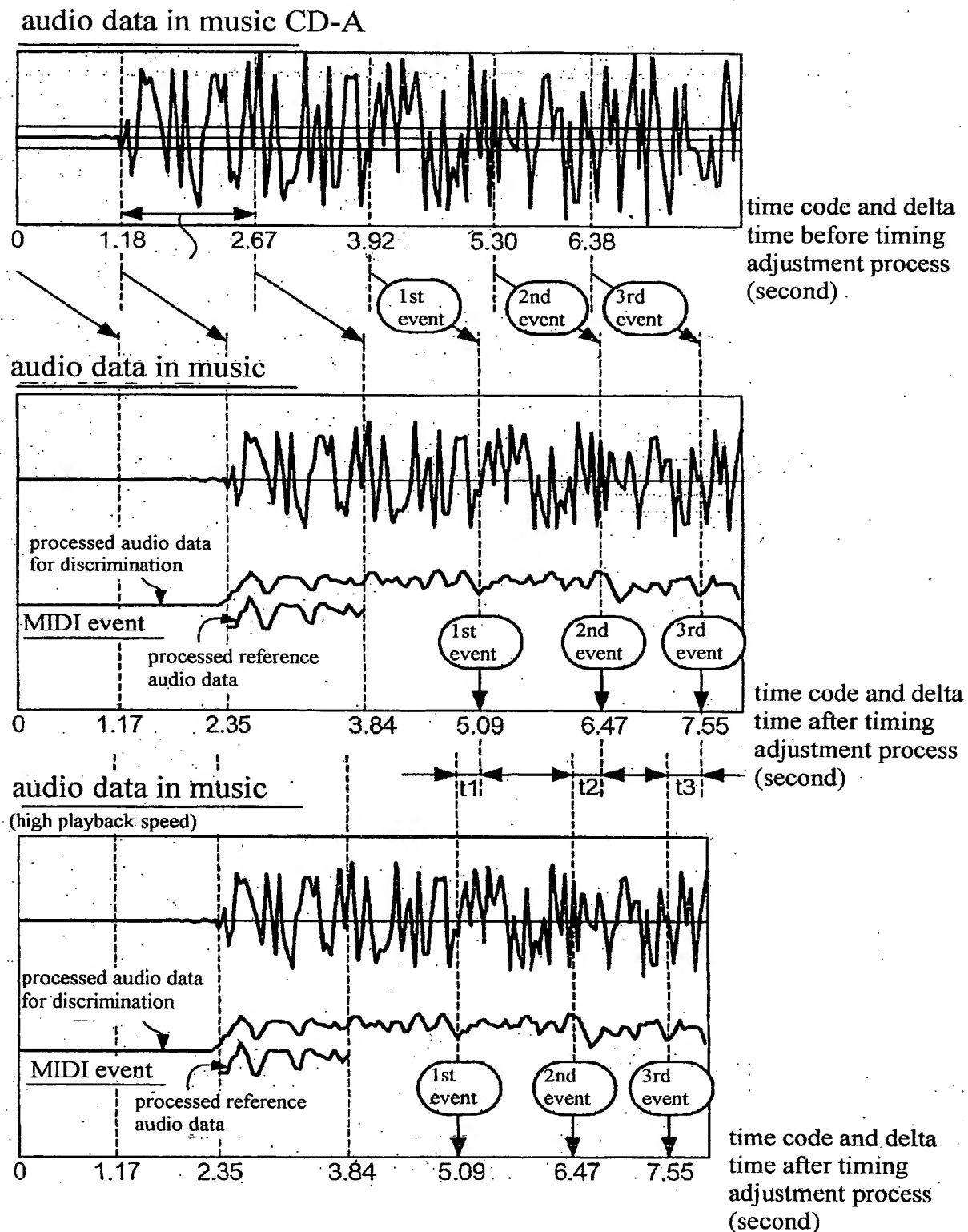
[Fig. 10]



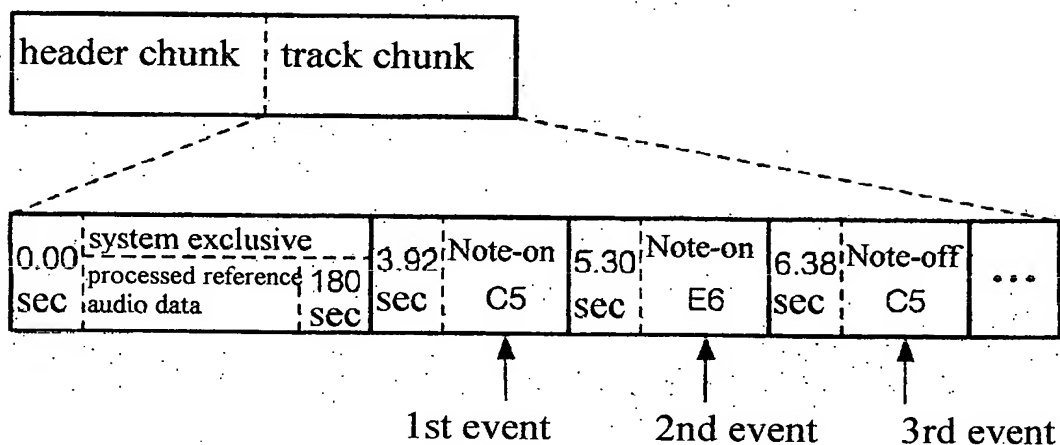
[Fig. 11]



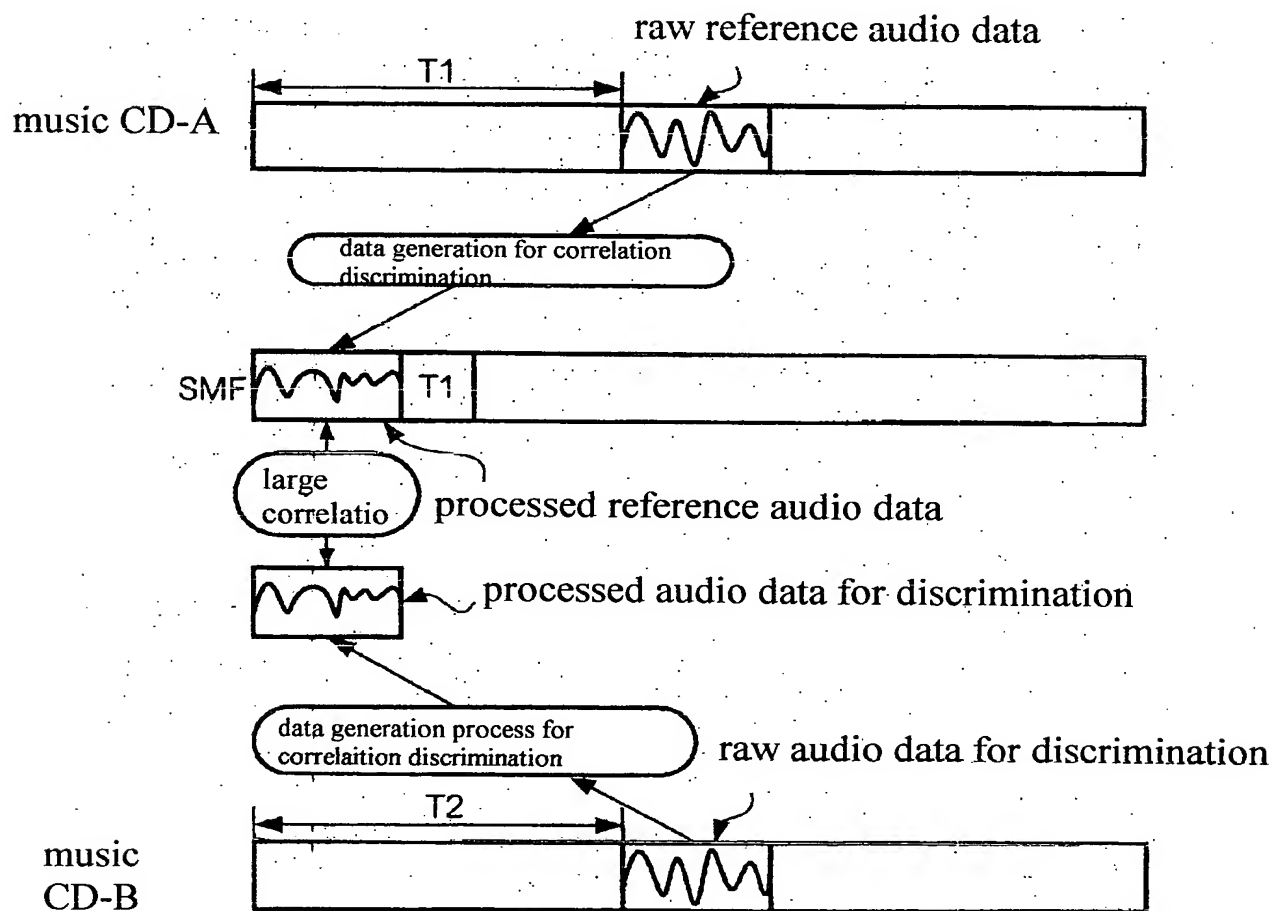
[Fig. 12]



[Fig. 13]



[Fig. 14]



[DOCUMENT NAME] ABSTRACT DOCUMENT

[ABSTRACT]

[PROBLEM] To provide a recorder, a player, a recording method, a playback method and program which permit synchronized playback of the performance data at correct timings for plural sets of audio data of the same musical tune having different starting timings.

[SOLVING MEANS] A controller 6 records MIDI data of performance through a piano 31, which is played with the playback of a music CD, in an SMF. At this time, the controller 6 generates reference audio data by using a part of the audio data in the music CD in order to record the reference audio data in the SMF. Subsequently, the controller 6 plays back the MIDI data recorded in the SMF together with the playback of the music CD. At this time, the controller 6 generates the audio data for discrimination by using the audio data in the music CD, and compares the reference audio data recorded in the SMF with the audio data for discrimination in order to determine the playback starting time of the MIDI data based on the comparison result.

[SELECTED FIGURE] Fig. 1